**Naval Research Laboratory**

**Washington, DC 20375-5320**

**NRL/MR/6180--08-9111**

# Intelligent Data Fusion for Wide-Area Assessment of UXO Contamination

## SERDP Project MM-1510
## 2007 Annual Report

Susan L. Rose-Pehrsson
Kevin Johnson

*Navy Technology Center for Safety and Survivability*
*Chemistry Division*

Christian Minor

*Nova Research, Inc.*
*Alexandria, VA*

Verner Guthrie

*U.S. Army Engineering Research and Development Center*
*Alexandria, VA*

February 29, 2008

# REPORT DOCUMENTATION PAGE

*Form Approved*
*OMB No. 0704-0188*



| **1. REPORT DATE** *(DD-MM-YYYY)* | **2. REPORT TYPE** | **3. DATES COVERED** *(From - To)* |
|---|---|---|
| 29-02-2008 | Annual | December 2006 – December 2007 |

**4. TITLE AND SUBTITLE**

Intelligent Data Fusion for Wide-Area Assessment of UXO Contamination

SERDP Project MM-1510
2007 Annual Report

**5a. CONTRACT NUMBER**
W74RDV53203308

**5b. GRANT NUMBER**

**5c. PROGRAM ELEMENT NUMBER**

**6. AUTHOR(S)**

Susan L. Rose-Pehrsson, Kevin Johnson, Christian Minor,* and Verner Guthrie**

**5d. PROJECT NUMBER**
SERDP Project MM-1510

**5e. TASK NUMBER**

**5f. WORK UNIT NUMBER**
61-9010-0-7

**7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)**

Naval Research Laboratory, Code 6180
4555 Overlook Avenue, SW
Washington, DC 20375-5320

**8. PERFORMING ORGANIZATION REPORT NUMBER**

NRL/MR/6180--08-9111

**9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)**

Strategic Environmental Research and Development Program (SERDP)
Dr. Anne Andrews/Dr. Jeffrey Marqusee
901 N. Stuart Street, Suite 303
Arlington, VA 22203-1853

**10. SPONSOR / MONITOR'S ACRONYM(S)**

SERDP

**11. SPONSOR / MONITOR'S REPORT NUMBER(S)**

**12. DISTRIBUTION / AVAILABILITY STATEMENT**

Approved for public release; distribution is unlimited.

**13. SUPPLEMENTARY NOTES**

*Nova Research, Inc., 1900 Elkin Street, Suite 230, Alexandria, VA 22308
**U.S. Army Engineering Research and Development Center, Topographic Engineering Center, Alexandria, VA 22315

**14. ABSTRACT**

Intelligent data fusion techniques are being developed and optimized for use in enhancing wide-area assessment (WAA) for UXO remediation efforts. This report describes the investigations in the second year of project MM-1510 that focused on data fusion. A generalized method for processing input data feature streams from UXO WAA surveys was developed. The method requires the generation of a corresponding geo-referenced feature intensity map and the specification of a functional relationship between the feature's intensity and the hypotheses supported by the presence or absence of that feature. The method accommodates diverse feature sets and widely varying evidential relationships as potential inputs for data fusion. Three data fusion theories were investigated: heuristic, Bayesian theoretic and Dempster-Shafer theoretic approaches. Two prototype data fusion frameworks were developed and evaluated with feature sets for the Pueblo and Kirtland sites. Preliminary results obtained with a prototype Dempster-Shafer based data fusion framework agreed well with the limited ground truth data available at the Pueblo site.

**15. SUBJECT TERMS**

Data fusion; Wide-area assessment; Unexploded ordnance; UXO; Remediation; Data management; Decision-making utility; Feature selection; Probability densities; Bayesian-based; Wavelets; Dempster-Shafer; Heuristic; Geo-referenced feature

| **16. SECURITY CLASSIFICATION OF:** | | | **17. LIMITATION OF ABSTRACT** | **18. NUMBER OF PAGES** | **19a. NAME OF RESPONSIBLE PERSON** |
|---|---|---|---|---|---|
| **a. REPORT** | **b. ABSTRACT** | **c. THIS PAGE** | | | Susan L. Rose-Pehrsson |
| Unclassified | Unclassified | Unclassified | UL | 72 | **19b. TELEPHONE NUMBER** *(include area code)* (202) 767-3138 |

**Standard Form 298 (Rev. 8-98)**
**Prescribed by ANSI Std. Z39.18**

# Table of Contents

# List of Acronyms

| | |
|---|---|
| AMTADS | Airborne Multi-Sensor Towed Array Detection System |
| EMI | Electromagnetic Induction |
| ESRI | Environmental Systems Research Institute |
| ESTCP | Environmental Security Technology Certification Program |
| GIS | Geographic Information System |
| GPR | Ground Penetrating Radar |
| GPS | Global Positioning System |
| LiDAR | Light Detection And Ranging |
| NRL | Naval Research Laboratory |
| nT | nanoTesla |
| PBR | Precision Bombing Range |
| SAR | Synthetic Aperture Radar |
| SERDP | Strategic Environmental Research and Development Program |
| TFM | Total Field Magnetometry |
| UTM | Universal Transverse Mercator Grid Coordinates |
| UXO | Unexploded Ordnance |
| WAA | Wide-Area Assessment |
| WAAPP | Wide-Area Assessment Pilot Program |

# List of Figures

# Acknowledgement

This research was supported wholly by the U.S. Department of Defense, through the Strategic Environmental Research and Development Program (SERDP). The authors thank the SERDP staff and team members for their assistance, particularly Dr. Herb Nelson and Dr. Dan Steinhurst.

# Executive Summary

**Background.** The remediation of sites contaminated with unexploded ordnance (UXO) remains an area of intense focus for the Department of Defense. Current estimates place the total area of possibly UXO-contaminated sites at 10 million acres, with an overall cost of remediation with current methods and sensing technologies in the tens of billions of dollars. Fortunately, studies have estimated that up to 80% of typical sites of potential contamination are actually UXO-free. What is needed to take advantage of this ratio is a means to quickly and reliably scan large sites (on the order of 10,000 acres) in order to rapidly identify regions that are free of UXO and regions that must be subjected to more detailed and time-intensive examination and remediation with established UXO detection tools. Recent investigations have focused on wide-area assessments (WAA) aimed at rapidly determining the approximate density and spatial distribution of UXO objects over regions of wide area, rather than identification of individual UXO objects. Several wide-area assessment projects have been completed under the auspices of Strategic Environmental Research and Development Program (SERDP) and Environmental Security Technology Certification Program (ESTCP). [1] These projects utilized various detection techniques, each with different strengths and weaknesses. However, no single sensing technology has been proven superior in wide-area assessment of UXO. It is therefore logical to examine data fusion approaches, which take advantage of all the available evidence, combining the strengths of each sensing technology while minimizing the weaknesses.

**Objective.** The objective of this work is to develop a data fusion framework that will form the basis of a cohesive data management and decision making utility for processing information acquired in the course of performing wide-area surveys of potential UXO remediation sites. This framework will be capable of capturing UXO-related information from all available data and effectively combining this information to provide site-wide assessments of the likelihood of UXO contamination that are more accurate than any single information source on its own. The final data fusion framework is intended to allow site managers to more efficiently direct the expenditure of time, labor and resources in remediation efforts.

This report details efforts undertaken in the second year of project MM-1510. The first year of this project examined potential wide-area assessment information streams, and determined the feasibility of feature selection methods for data fusion. The second year of the project focused on the development of an architecture and data fusion algorithms for the framework. Future work will focus on optimization and refinement of the methodology and algorithms with the development of a prototype implementation of the data fusion framework and its evaluation at several sites.

**Theory.** In the context of UXO remediation, intelligent data fusion is the combination of unbiased pattern recognition techniques with expert information about the strengths and weaknesses associated with the data acquisition techniques and expert knowledge of the geology, foliage, and man-made features located at the site under scrutiny. Data from

Manuscript approved January 18, 2008.

multiple sensing modalities are utilized in an overall decision-making algorithm that is theoretically more accurate than any individual sensor on its own. [2-4]

The two most well-known theoretical data fusion approaches are those formulated by Bayesian inference theory and by Dempster-Shafer theory. Bayesian theory, the older and more established of the two, is founded on a rigorous statistical framework and generally requires significant statistical assumptions regarding the input data. Dempster-Shafer can be thought of as a relaxation of Bayesian theory in which subject belief assessments are utilized and can be assigned to supersets of hypotheses as well as to single hypotheses. In this work, UXO WAA data-specific approaches based on both theories were implemented and evaluated, along with a simple heuristic approach representing current geographic information system (GIS) techniques relying on data overlay and visualization.

Site assessment with the data fusion framework proceeds as follows: multiple data sets utilizing different wide-area sensing technologies are acquired at a site potentially in need of UXO remediation. Additional information about the site, including locations of potential UXO contamination and relevant historical, geological, and topological data, are also acquired. All data sets are geo-referenced, processed, and/or analyzed in some fashion to extract UXO-related features. Each feature set is converted into a site-wide feature map, referred to as a feature layer, which serves as an input into the data fusion framework. Features are associated with UXO contamination in a quantitative fashion, and the input feature maps are transformed into assessments of the presence or absence of UXO, conditioned on evidence provided by that feature. These quantitative assessments are combined on a point-by-point basis throughout the survey area via a data fusion algorithm to provide a site-wide, overall assessment of UXO contamination. The process is inductive by design. New information and more refined data sets can be incorporated as they become available.

**Results.** The major accomplishments for Year Two are listed below:

- A wavelet filtering algorithm was developed to improve the automatic crater identification algorithm. Filtering reduced false positive crater identifications by removing surface texture artifacts.
- A statistical nearest-neighborhood clustering filter algorithm was developed to improve the magnetic anomaly data by locating anomalies that were part of non-uniform groupings. The clustering algorithm reduced background noise and false positive assessments of UXO in the final data fusion output.
- A Data Fusion Framework prototype based on a hybrid Dempster-Shafer theory was developed.
- A methodology for the input and registration of disparate data and feature streams from wide-area sensing technologies and other site-specific knowledge was devised.

- A methodology allowing the input of meta-information regarding the relationship between these lines of evidence and the presence (or absence) of UXO or UXO-related objects was implemented.
- A method by which heuristic rules can be incorporated into the Data Fusion Framework was developed in order to take advantage of specific known interdependencies between feature layers.
- Demonstration of the Data Fusion Framework prototype's ability to successfully provide useful output assessment of UXO likelihood from an assortment of WAA assessment data and features.

The principal accomplishment of Project MM-1510 was the development of a prototype Data Fusion Framework suitable for wide-area assessment of UXO contamination. A key enabling technology was the development of a generalized method for processing input data feature streams from UXO WAA survey efforts. The development of this method is significant, as it requires only a limited number of specifications to be imposed on input features, allowing a wide range of feature sets and relationships to be formatted and input for data fusion. Such flexibility is crucial, as the disparate nature of the data and features available from potential WAA survey techniques presents a significant impediment towards adoption of more basic data fusion approaches.

Project MM-1510 leveraged data and feature sets acquired in the ESTCP Wide-Area Assessment Pilot Program (WAAPP). Various feature sets derived from both ESTCP performers and from customized feature extraction algorithms developed at NRL were utilized as inputs to data fusion, demonstrating the flexibility of the approach. Successful feature extraction algorithms developed in year one were further optimized and validated against new data, including an automatic crater detection algorithm and a nearest-neighbor clustering filter to selectively remove false-positive magnetometry anomalies. Methods for generating feature layers from extracted features for input to data fusion were developed and implemented. For each input feature set, a corresponding feature intensity map and specification of a functional relationship between a feature's intensity and the likelihoods for or against the presence of UXO that are supported by the feature's intensity values. Heuristic, Bayesian, and Dempster-Shafer theoretic algorithms for combining evidence presented in feature layers were investigated as possible engines for a UXO WAA data fusion framework prototype. These were implemented as MATLAB code and evaluated with feature layers generated from both the Pueblo and Kirtland site data acquired by performers in the ESTCP WAAPP.

The Dempster-Shafer approach, with its ability to quantify uncertainty about evidence, was shown to be the most appropriate data fusion strategy for the UXO problem and proved to be the most successful of the three. The ability to incorporate heuristic rules regarding specific dependencies between input feature layers into the Dempster-Shafer based data fusion framework prototype was described and demonstrated utilizing two specific examples. The first demonstrated a reduction of false positive indications of UXO by utilizing a feature layer comprised of manually identified man-made structures to selectively block magnetometry-derived features. The second demonstrated an

adjustment of the impact of magnetometry-derived features on the output assessment of UXO to accurately reflect the uncertainty associated with increased magnetometry data sparseness in some areas of the helimag survey. The prototype data fusion framework developed was able to delineate areas of likely contamination while providing reasonable estimates of the likelihood of that contamination given supporting observational evidence and a priori knowledge. Preliminary results were compared with limited ground truth data available at the Pueblo site and agreed well.

**Benefits.** The key theoretical advantage of a data fusion approach to wide-area assessment is the ability to reduce false positives while retaining high detection rates. The framework described is flexible, tolerating missing data and allowing multiple configurations of potential input data streams, as well as scalable, allowing new data streams to easily be included in the assessment. Further, the impact of available and new data streams on the output can be readily quantified. One challenge is that the structured input methodology requires the specification of each feature layer's relationship to the presence or absence of UXO. However, the input methodology allows specification to be accomplished in a highly flexible manner. The user has the ability to input specifications that vary from simple, intuitive estimations based on expert knowledge to detailed functional relationships based on empirical evidence of sensor performance. Thus, the data fusion framework is capable of utilizing all the information and observation evidence available, without necessarily requiring that the exact same inputs be present for assessment. This flexibility is an important feature of the data fusion approach as it is expected that, for a number of reasons, it will rarely be the case that exactly the same types or quality of data will be available for analysis each time a wide-area UXO assessment is performed.

**Future Applications.** MATLAB was chosen as the computational platform for data exploration and initial feature and algorithm development. [5] An extremely useful research tool, MATLAB provides an extensive code base of algorithmic resources together with the ability to rapidly port prototype implementations to other computational platforms. In the final year, a prototype software implementation of the data fusion framework will be developed that is suitable for demonstration and evaluation by SERDP at various sites. In addition, the development of a final production-grade data fusion framework that is well-suited for independent use by site administrators will be planned in close coordination with SERDP and potential end-users, and in partnership with an existing GIS software vendor. It is expected that the development and demonstration of the final production-grade data fusion framework will be completed under a follow-on ESTCP program.

# Objective

The objective of project MM-1510 is to enhance assessment of unexploded buried ordnance (UXO) in large geographic areas through data fusion of outputs of multiple sensing technologies with any available expert knowledge of the sites in order to reduce overall false alarm rates. Project MM-1510 centers on the development of a cohesive data management and decision making strategy for analyzing UXO survey data and site-specific information. The development of such a data fusion framework requires the development of a customized algorithmic data fusion approach that enables the effective combination of all available knowledge about a UXO survey site (including both data feature analysis and manually-acquired site information) into output wide-area maps of potentially contaminated sites delineating areas of likely UXO contamination from those that are likely to be free of UXO. The intention is that these assessments will serve to direct the acquisition of data using more accurate, sensitive, and costly local-area UXO surveys. Data from local-area surveys can then be fed back into the data fusion framework and used to further refine maps of these regions and to provide more detailed assessments and allowing more efficient direction of time, labor, and resources.

This report describes the second year of a three-year effort to develop the data fusion framework. The overall goal for the second year of project MM-1510 was the development of appropriate architecture and algorithm components of the data fusion framework for enhancing wide-area assessment of UXO. Towards this end, data fusion strategies and algorithms were explored and developed, feature layer development efforts begun in year one of the project were continued with an eye towards ensuring compatibility with and suitability for potential data fusion frameworks, and a prototype implementation of the developed data fusion framework was created to test and evaluate output assessments.

# Background

### Problem Statement

Current estimates place the total area of possibly UXO-contaminated sites at 10 million acres, with an overall cost of remediation with current methods and sensing technologies in the tens of billions of dollars. Studies have estimated that up to 80% of typical sites of potential contamination are actually UXO-free, indicating a need to quickly and reliably scan large sites (on the order of 10,000 acres) in order to rapidly identify regions that are free of UXO and regions that must be subjected to more detailed and time-intensive examination and remediation with established UXO detection tools.

Effective wide-area UXO assessment is centered on the capability to rapidly scan large tracts of land and obtain relevant, useful information in the process. Two possible modes of enhancement to wide-area UXO assessment are, first, decreasing the false alarm rate

of current UXO sensing technologies, and second, utilizing alternate sensing technologies and survey methods that scan larger areas or more rapidly cover large areas than current sensing platforms.

## Current Technology

Standard, ground-based UXO sensing technologies include methods such as vehicular-mounted time and frequency domain electromagnetic induction (EMI), total field magnetometry (TFM), and ground penetrating radar (GPR).[*] These UXO sensors have been deployed on ground-based platforms such as portable devices, push carts, and towed sensor arrays. Standard analysis methods of these types of data are well described in the literature and have been implemented with success. These methods typically rely on generating theoretical sensor response models, or measuring pure responses of various UXO items and then comparing survey data to these models in order to make a detection or classification. Typically, the ability of such techniques to cover wide survey areas is limited, although current SERDP projects are assessing the utility of ground survey transects to increase this capability for wide-area applications.

As the sensitivity/ ground penetration depth of direct UXO sensing technologies, such as magnetometry, drops steeply with distance, alternate techniques for wide-area assessment will generally not have the benefit of sensing deeply buried UXO objects directly. Instead, they must rely instead on sensing UXO-related phenomena like spectral chemical signatures, variations in heat capacity, and measurements of surface clutter and micro-topological features. The techniques of synthetic aperture radar (SAR), light detection and ranging (LiDAR), and high-resolution aerial photography all yield information about micro-topological features and surface clutter of wide areas under assessment. These sensors have the benefit of functioning over much greater distances than electromagnetic sensors, allowing them to be deployed on fixed wing aircraft. Towards this end, six wide-area assessment projects have been completed under the auspices of SERDP and ESTCP. [1] These projects utilized techniques that varied from airborne infrared laser imaging combined with thermal imaging (UX-9523), synthetic aperture radar (UX-0126, UX-1070, and UX-1173), and airborne magnetometer arrays (UX-0031 and UX-3002). Additionally, ESTCP has sponsored a "Wide-Area Assessment Pilot Program" to fund development of wide-area UXO sensing technologies. [6]

Wide-area sensing platforms are typically designed and optimized to detect UXO or UXO-related features without correlation to other sources of data or expert information. Data are acquired and processed via algorithmic feature extraction, manual inspection and feature identification, or some combination of both. Final site assessment is provided by evaluating the extracted UXO-related features, taking into account a priori knowledge as to how accurate the features are at indicating the presence of UXO and/or UXO-related objects. Thus, two types of information are obtained from sensing: UXO-related features extracted from survey data, and meta-data regarding the relationship between

[*] See "*Proceedings of SPIE:* Detection and Remediation Technologies for Mines and Minelike Targets V through VIII" and "*Transactions on Geoscience and Remote Sensing, Vol 2, No. 3, (2001)*" for examples.

UXO and these features. Further information can be obtained from manual assessment of the site. Areas associated with UXO-related activities can be delineated through visual inspection of sensor data or through knowledge of historical usage patterns at the site. Examples of this type of information include delineations of visible bombing targets and descriptions of known munitions ranges.

Thus, at each survey site, site managers potentially have multiple, disparate lines of evidence for or against the presence of UXO or UXO-related items. The goal of data fusion is to produce the best assessment of UXO contamination at any location in the site, given all the available evidence, and in turn, allowing site managers to more accurately delineate areas of likely UXO contamination.

**Data Fusion Approach**

To date, effective wide-area assessment has been hindered by a lack of accurate target and range information. In addition, no single sensing technology has been both accurate and cost-effective in surveying entire sites. One potential avenue of enhancing wide area assessments of UXO is to combine multiple data streams, taking advantage of the strengths of individual sensing platforms, while minimizing the weaknesses. [4]

Limited data fusion approaches have been explored, although these tend to be data-stream specific rather than generic data aggregators and decision utilities. [7-10] The approach undertaken in this effort is to enhance wide area assessment of UXO by developing an algorithmic framework for data management and data fusion that provides a structured, intuitive means for incorporating and inputting UXO WAA survey sensor data, algorithmically-generated data features, manually located features, site-specific information, and any relevant associated meta-data, such as sensor performance characteristics and specific heuristic rules regarding data stream interdependencies.

The central task in this effort is to find a structured way of combining these multiple lines of evidence, and information. The nature of the wide-area UXO assessment problem suggests three main criteria for any proposed data fusion framework. The first criterion is that a potential data fusion framework must be capable of accepting a variable number of disparate data streams as input. Different types of information are conveyed by each stream, and not every line of data will be available at every geographic grid point in the survey area. Thus, the data fusion framework must be capable of producing output assessment that utilize the data available, rather than a fixed set of data inputs.

The second criterion is that a potential data fusion framework has a structured, intuitive means of capturing the relationship between a given data feature or other form of information and the likelihood of the presence of UXO. The reality of the UXO WAA problem is that extensive ground truth data will likely never be available, due to the considerable expense required to acquire it. This complicates efforts at statistically modeling the relationship between extracted features and UXO contamination and makes precise probabilistic assessments of UXO contamination highly difficult to achieve. Thus, it makes sense to focus on approaches that allow the user to incorporate and benefit

from knowledge that is known regarding the specific performance characteristics of the individual sensing technologies, as well as that which can be gleaned from the expertise of seasoned site managers and other program personnel.

Further, if expert information is known regarding the performance characteristics of one of the wide area assessment technologies in detecting UXO, then the data fusion framework should be capable of directly accepting and utilizing this information. Likewise, if the relationship between an information stream and UXO is not well-known, then there should be a structured, intuitive means of inputting reasonable estimates of this relationship into the data fusion framework. These relationships should not be "hard wired" into the framework itself, as this would not allow for updating in the face of improved knowledge of sensor characteristics, or in site-specific conditions that may alter such relationships from site to site. Finally, the data fusion framework should be able to incorporate these relationships independently of their functional form to reflect the fact that, in some cases, complex functional dependencies may be know from empirical testing while in others, only simple estimates arising from the site manager's personal experience may be available.

Finally, the third criterion is that a data fusion framework must be reasonably amenable to the incorporation of specific heuristic rules that would otherwise be difficult to capture in a strictly probabilistic fashion. These rules may reflect specific, known interdependencies between two or more different input data streams and would operate on the input information to adjust the output assessments accordingly.

# Methods

In the broadest terms, data fusion for wide area assessment of UXO proceeds as follows: multiple data sets utilizing different wide-area sensing technologies are acquired at a site potentially in need of UXO remediation. Additional information about the site, including locations of potential UXO contamination and relevant historical, geological, and topological data, are also acquired. All data sets are geo-referenced, processed, and/or analyzed in some fashion to extract UXO-related features. Each feature set is converted into a site-wide feature map which is input into the data fusion framework. Each type of feature is associated with UXO contamination in a quantitative fashion, and the input feature maps are transformed into assessments of the presence or absence of UXO, conditioned on evidence provided by that feature. These quantitative assessments are combined on a point-by-point basis throughout the survey area via a data fusion algorithm to provide an overall assessment of UXO contamination. The process is inductive by design. New information and more refined data sets can be incorporated as they become available.

Fundamental to the data fusion approach is the development of a standardized input method for sensor data, features, and expert information. The specification of a common format for features from varying sites was the focus of year one of project MM-1510 and has continued to figure importantly in work done in year two. The approaches pursued are described in the section on feature layer development that follows.

With the inputs established, the development of a data fusion architecture began with the formulation of an appropriate frame of discernment for the problem of wide-area assessment of UXO. Several theoretical data fusion approaches were evaluated for their suitability. Methods for performing data fusion with heuristic, Bayesian theoretic and Dempster-Shafer theoretic approaches are presented below. Finally, a prototype implementation of a data fusion framework embodying the identified data fusion strategies was constructed in software in order to evaluate the strengths and weaknesses of the approaches and to examine possible avenues for optimization.

Project MM-1510 leveraged data acquired by SERDP/ESTCP programs, and in particular, data acquired for three former Department of Defense sites during the ESTCP Wide-Area Assessment Pilot Program. The three sites surveyed in 2004 and 2005 were: the Pueblo Precision Bombing Range #2 in Colorado (CO), the former Kirtland Bombing Targets N1 and N3 in New Mexico (NM), and the Military Wash Area in the Borrego Maneuver Area in California (CA). The principle sensing modalities utilized in each survey were low altitude (helicopter) airborne magnetometry, often referred to as "helimag," high-altitude airborne LiDAR, and orthophotography. Along with sensor data, a number of extracted features were made available. These included manually identified regions of interest from LiDAR and photographic data and magnetic anomalies from helimag data. Additional feature sets were generated algorithmically in the first year of project MM-1510, including automatically identified craters and magnetic

signatures. [11] Although the work described in this report was applied to these data sets, it should be noted that the underlying approach developed for assimilating data is intended to apply to any UXO-related feature data that are available for inclusion in the data fusion framework.

**Feature Layer Development**

In general, UXO-related features are extracted from acquired data either manually or through application of a feature extraction algorithm. These features fall into one of three categories of feature type: features with continuously-variable intensity, binary features that are point located, and binary features resulting from the delineation of regions of interest. For example, potential crater locations can be extracted from LiDAR and aerial photography through either visual inspection or automated crater detection algorithms. These features are binary in nature and point-located. Other extracted features could be continuously valued parameters, such as background-corrected total field magnetometry signals extracted from magnetometry data. In this case, the extracted features have both location and intensity. Finally, manual assessment of the site, either through visual inspection of the data or through knowledge of historical usage patterns at the site, can result in delineated areas that are associated with UXO-related activities. Examples of this type of data include delineations of visible bombing targets and known munitions ranges.

Extracted features constitute multiple, disparate lines of evidence for or against the presence of UXO or UXO-related items at each survey site. While each feature can be visually displayed, and even overlaid on a site map, the disparity between the different feature sets makes it difficult to algorithmically combine the evidence provided by them into an overall assessment. In order to overcome this difficulty, a data fusion architecture was developed in which individual feature types were used to produce corresponding feature "layers." In this approach, the information encapsulated by each available extracted feature was converted into a map indicating the strength of that feature across the survey site. This allowed for a pixel-by-pixel determination of feature specific evidence related to UXO contamination throughout the survey site. The salient property of each feature layer was that it conveyed a site-wide map filled with feature intensity values ranging between a known minimum and maximum value. Survey grid points where no feature information was available were allowed and were treated as missing data or estimated via interpolation from nearby data. Along with a structured representation of each layer's relationship to UXO contamination, these feature layers were utilized as inputs into subsequent data fusion algorithms responsible for synthesizing the evidence presented by each feature layer into a site-wide, point-by-point assessment of the likelihood of UXO contamination.

Generation of individual feature layers was accomplished within the MATLAB computational platform. [5] Feature layers were organized in the same file structure as that used to store imported survey data. [11] The method by which each feature layer was generated depended entirely on the nature of the extracted feature. Continuously-valued features, such as magnetometry survey data, were used directly as feature layers.

Manually delineated feature areas, such as visually identified bombing targets or historically known munitions areas, were converted into a binary feature layer in which grid points within the features were given a value of one and points outside a value of zero. Delineated feature areas for bombing targets were further dilated to encompass an "effective area" that more closely resembled the likely area of UXO contamination from repeated bombing runs. Collections of point-location binary features, such as the locations of identified craters or magnetic anomalies, were transformed into site-wide feature density estimates via a kernel density estimation (KDE) algorithm. [12] For simplicity, the generated feature density maps were scaled according to their maximum value, and thus reflected the relative density of the feature across the site.

By utilizing a feature density estimate, the likelihood of UXO contamination was correlated to the intensity of feature activity in a nearby region, rather than to the presence or absence of the feature itself at any specific point. The size of the region of influence was chosen to roughly correspond to the effective area of a bombing target by doubling the approximate length of the major axis of each target and replacing the target feature with a circle of that diameter, centered upon the bombing target's location. In kernel density estimation, points most distant from the center point being estimated are scaled to make the least contribution, nearby points the most contribution. Gaussian scaling was used for all KDE-generated feature density maps.

**Data fusion architecture development**

The next task was to develop a data fusion methodology to accurately assess UXO contamination given the evidence provided by an arbitrary set of feature layers. Three algorithmic approaches were considered for implementation: a purely heuristic-based linear combination, a Bayesian theoretic approach, and a Dempster-Shafer theoretic approach. In addition to developing an algorithmic approach for combining the evidence presented by multiple feature layers, a standardized, structured means for defining the relationship between each feature layer and the presence of UXO was required.

The simplest approach involved the use of heuristic rules to weight each feature layer according to its contribution to the evidence that UXO was present and then summing the contributions to form an output assessment. Intuitively, a heuristic approach is equivalent to overlaying each available feature layer in a visual display, and then adjusting the opacity of each layer according to its importance in signifying UXO contamination. Specific heuristic rules can come from expert knowledge regarding the features and their relationship to UXO, or through analysis of empirical data to discern relationships. These rules express UXO likelihood as a function of feature layer value (i.e., intensity) that varies between zero and one, depending on the degree to which that feature layer expresses support for the presence of UXO. While a heuristic approach is simple computationally and allows for detailed input of specific UXO-feature relationships, a number of problems present themselves in implementation. Chief among these is the lack of a measure of uncertainty or probability in the output of such an algorithm, and no means for incorporating such information as an input. This lack puts reliable information streams on par with less reliable lines of evidence in terms of its impact on the final

output assessment. Additionally, experience with fusion of multiple sensor outputs at NRL has shown that a simple "AND" operation such as this often provide little benefit as they typically increase false alarms along with detections. For this reason, the purely heuristic approach was eliminated from further consideration.

Next, a Bayesian theoretic approach was considered. Bayesian reasoning and Bayesian belief network architectures are powerful tools for inductive reasoning within a probabilistic framework. [13,14] The core of the approach was formulated from Bayes' rule which states that for a hypothesis, *H*, and evidence, *E*, the probability of the *H*, conditioned on *E* is:

$$p(H|E) = p(H) \times \frac{p(E|H)}{p(E)}$$

The first term, *p(H)*, is known as the prior probability of *H* and represents knowledge of *H* before observing evidence *E*. The second term is known as the normalized likelihood and represents knowledge of how likely it is to observe *E*, given *H*. Multiplication of these two terms provides the posterior probability of *H*, given observed evidence *E*. If multiple lines of evidence are available, the joint distribution must be considered:

$$p(H|e1, e2, e3,\ldots) = \frac{p(H,e1,e2,e3,\ldots)}{p(H,e1,e2,e3\ldots) + p(\neg H,e1,e2,e3\ldots)}$$

Fortunately, in many cases calculating the joint distribution of several variables is made simpler by examining the causal relationships among the variables and identifying those which are independent. A specification of these relationships and the conditional probability values associated with them forms what is known as a Bayesian belief network.

A diverse range of applications have been described for belief networks, such as automated medical diagnosis and intelligent software help utilities. Of particular utility is the ability of such networks to optimally estimate unknown network parameters from known parameters, allowing, for example, diagnosis of a particular disease with a limited set of observations and test results. This ability, however, stems from the belief network's complete encapsulation of all pertinent probabilistic relationships between network parameters. As such, belief networks are only useful when an appropriate network topology is defined with appropriate causal links such that the resulting conditional probability distributions assigned to each link are fully specified. This is problematic for the UXO WAA task, as estimating either of these quantities can be difficult, and possibly site-specific. In other words, it's possible that not only will the types of evidence available change from site to site and, indeed, from location to location within a given site, but also that the conditional probability distributions and network topology itself will also change.

A simpler, so-called "naïve" Bayesian theoretic approach is to assume independence amongst each line of evidence. Beginning with an initial prior probability for UXO, the

prior can be updated to a posterior probability of UXO by multiplying it with the normalized likelihood of each available line of evidence.

$$p(UXO|E) = p(UXO) \times \prod_{i}^{n} \frac{p(e_i|UXO)}{p(UXO)\,p(e_i|UXO) + p(\neg UXO)p(e_i|\neg UXO)}$$

where $E$ is the set of available lines of evidence:

$$E = \{e_1, e_2, e_3, \ldots e_n\}$$

Utilizing this formulation, a naïve Bayes data fusion algorithm was implemented within MATLAB. In this implementation, conditional probability functions for $p(e|UXO)$ and $p(e|\neg UXO)$ must be defined across the range of observed feature intensity values for each feature data layer, representing the probability of observing a particular feature value, given that UXO is present and not present, respectively. These functions were defined a priori and were estimated from knowledge regarding the feature's relationship to UXO. To begin the data fusion process, an uninformed prior probability of UXO in the absence of any evidence was assumed, and was represented by setting $p(UXO)$ to an initial value of 0.5 at each point in the survey grid. For each input feature layer, $p(UXO)$ was updated at each point in the survey grid to a posterior probability of UXO, $p(UXO|E)$, given the evidence observed in that feature layer. The specified conditional probability functions were used to convert feature intensity values into conditional probability assignments on a point-by-point basis across the survey grid. These assignments were then used to calculate the normalized likelihood ratio at every point in the survey grid and then multiplied by the corresponding prior probabilities and generating a posterior probability for every grid point. These posterior probabilities were then updated with subsequent feature layers in the same fashion, until all available feature layers were included. The resulting output was a single site-wide map containing values for the posterior probability that UXO was present, given the observed feature data layers. Thus, values close to one indicated a high likelihood that UXO was present, those close to zero indicated little likelihood that UXO was present, and those close to a value of 0.5 indicated relative uncertainty as to whether UXO was or was not present.

The main difficulty in implementing a Bayesian theoretic approach is that it requires specification of the conditional probabilities $p(e|UXO)$ and $p(e|\neg UXO)$, which can be difficult to do in an intuitive, rational manner depending on the nature of the feature layer and its relationship to UXO. A related problem is that within this formalism, a lack of evidence for a supposition is generally interpreted as evidence for that supposition's negation (i.e., $p(x) + p(\neg x) = 1$). For example, if a feature layer represents a direct detection of UXO-related objects, then a specific feature layer value can be considered as a detection threshold and $p(e|UXO)$ can be related to the feature's UXO detection rate at that threshold, and $p(e|\neg UXO)$ can be related to the feature's false positive rate at that same threshold. However, other feature layers may encode a relationship that is difficult or even nonsensical to parse in this manner. For example, crater density may be indirectly associated with UXO-related objects, but cannot be viewed as a detection mechanism in the classic signal detection framework, as it would involve assigning a

probability to particular crater density conditioned on the fact that UXO is or isn't present. Unfortunately, these probability assignments are tenuous at best due to the lack of a causal mechanism between the two features. For the UXO problem, the fundamental difference between the two lines of evidence, $p(e|UXO)$ and $p(e|\neg UXO)$, is the type of uncertainty associated with them. The former (i.e., $p(e|UXO)$ that links crater density with UXO-related objects) is associated with objective uncertainty and is well-described by traditional probability theory. The latter is associated with subjective uncertainty in so much that a lack of craters reflects ignorance towards either the presence or absence of UXO contamination.

In order to avoid the complications encountered in probability assignments for the Bayesian method, an approach based on Dempster-Shafer theory was also developed. Briefly, Dempster-Shafer theory can be described as a generalization of Bayes theory in which observed evidence can support not only specific hypotheses, but also sets of hypotheses. [15,16] The approach is attractive for problems involving evidence pooling from multiple sources because it does not require complete specification of the underlying conditional probabilities and it allows for assignment of a degree of belief to a specific hypothesis without necessarily assigning any belief to the negation of that hypothesis. These properties enable Dempster-Shafer frameworks to address evidence associated with subjective uncertainty in a more satisfying manner while simultaneously retaining the ability to incorporate evidence associated with objective uncertainty. [17] Dempster-Shafer data fusion has been utilized in applications such as land cover classification, machine vision, and medical diagnoses. [18-20]

An example of Dempster-Shafer theory applied to sensors follows: consider a hypothetical sensor system that attempts to identify a sensed object as belonging to one of four object types: A, B, C, or D. The exhaustive and mutually exclusive hypothesis space, or frame of discernment, is represented by the set {A,B,C,D}. The evidence observed from a given sensor response provides support for one or more of these elements including, potentially, supersets combining two or more elements. Assignment is accomplished through the apportionment of unit probability mass across the focal elements for which the sensor's response provides support. The amount of mass assigned to any given focal element is representative of the relative amount of certainty with which the sensor response can make that declaration. For example, say the output of sensor *1* indicates that the sensed object is of type A or B with 80% certainty. Thus, an assignment of 0.8 is made to the hypothesis that the object is type A or B, $m_1(A,B) = 0.8$. The remaining mass of 0.2 can be assigned to this hypothesis's negation, namely that the object is type C or D, $m_1(C,D) = 0.2$. Alternatively, it can be assigned to superset of all elements in the frame of discernment, reflecting total uncertainty as to the object's identification, $m_1(A,B,C,D) = 0.2$. The former assignment is useful in situations where the evidence is well-described by traditional probability assessments, while the latter assignment is useful in describing evidence associated with subjective uncertainty or incomplete probability specifications.

Probability mass assignments from two sensors are pooled according to Dempster's rule of combination. Pooling is accomplished by calculating the cross products of the

probability mass assignments for each sensor. Each cross product is assigned to a hypothesis that represents the intersection of those of the two component probability masses used to generate it. Cross products assigned to the same hypothesis are summed, resulting in a series of unique output hypotheses with corresponding probability mass assignments. If any cross products result from conflicting hypotheses, they are removed from the set and the remaining masses are normalized to sum to unit mass. Such renormalization has the effect of redistributing the mass assigned to conflict proportionally across the remaining hypotheses. For example, combining the evidence provided by the sensor in the previous example with that of a second sensor with probability mass assignments, $m_2$(A,C) = 0.7, $m_2$(A,B,C,D) = 0.3, proceeds as:

| | | $m_2(A,C)=0.7$ | $m_2(A,B,C,D)=0.3$ |
|---|---|---|---|
| *Sensor 1* | $m_1(A,B)=0.8$ | $m_{1,2}(A)=0.56$ | $m_{1,2}(A,B)=0.24$ |
| | $m_1(A,B,C,D)=0.2$ | $m_{1,2}(A,C)=0.14$ | $m_{1,2}(A,B,C,D)=0.06$ |

*Sensor 2*

Here, the net effect of data fusion is the assignment of most of the probability mass to $m$(A) (=0.56), a reduction of those assigned to $m$(A,B) (=0.24) and $m$(A,C) (=0.14), as well as a reduction of the mass assigned to uncertainty $m$(A,B,C,D) (=0.06).

Finally, the quantities of support and plausibility for each focal element are central to Dempster-Shafer theory. Support is defined as the sum of probability masses that can be directly attributed to that focal element. Plausibility is defined as the sum of probability masses not assigned to the focal element's negation. Together, the two quantities form an uncertainty interval bounded on the lower side by the support and on the upper side by the plausibility. The uncertainty interval conveys information regarding the proportion of evidence that directly supports a focal element versus that which merely fails to negate it. For example, support for A in the example above is 0.56 following fusion, while the plausibility is 1. This indicates that A receives partial support from the evidence, while no evidence directly refutes it. Thus, an uncertainty interval of (0,1) indicates complete uncertainty, intervals of (0,0) and (1,1) indicate focal elements known to be false and true, respectively, and intervals in which the support and plausibility are equal indicate that all of the evidence available directly supports that focal element.

A customized Dempster-Shafer data fusion algorithm was implemented within MATLAB. For this implementation, a frame of discernment of {*UXO*, ¬*UXO*} was used. Thus, there were three separate focal elements to which probability assignments can be made by feature-based evidence: {*UXO*}, reflecting the belief that UXO was present, {¬*UXO*}, reflecting belief that UXO was not present, and {*UXO*, ¬*UXO*}, reflecting ignorance (i.e., uncertainty) regarding the presence of UXO.

As with the Bayesian approach, the heart of the implementation involved the input of a series of feature layers in the form of site-wide maps of feature intensity values. Each feature layer was accompanied by a corresponding probability mass assignment function that related the probability mass assignments for each focal element to feature intensity over the range of feature intensities observed.

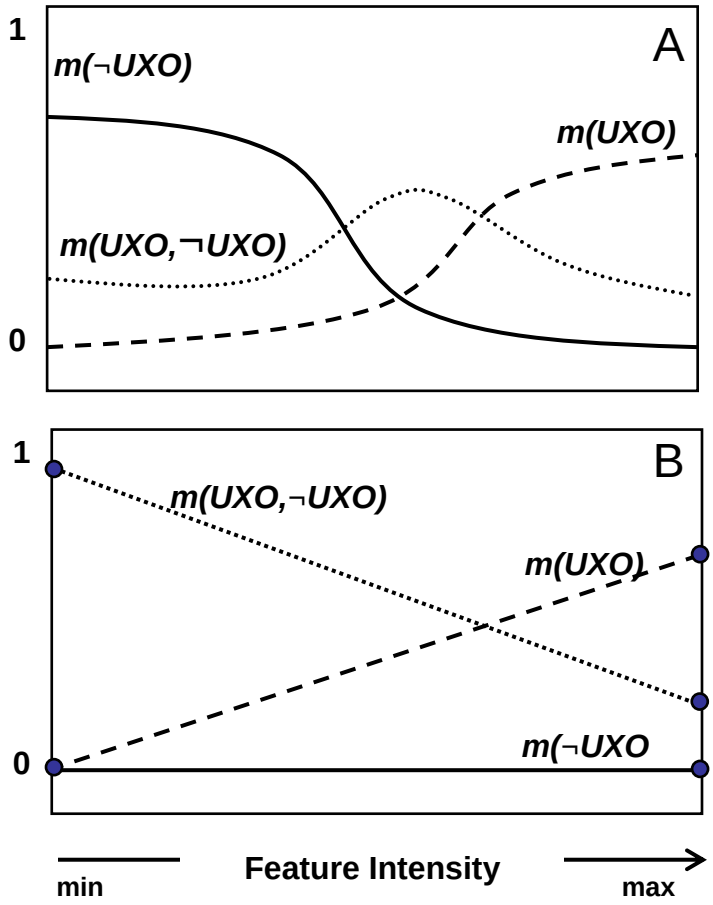



Figure 1. Hypothetical probability mass assignments in Dempster-Shafer data fusion. Example A depicts the assignment of complex functions to each of the three focal elements over the observed range of feature intensity values. Example B depicts an assignment where probability masses are assigned to the feature intensity extremes for each of the three focal elements and intermediate values are estimated through linear interpolation.

These functions were assigned a priori and were based on expert knowledge of the feature and its relation to the presence of UXO. If only limited knowledge is available, simpler estimations can be made. Two hypothetical examples are shown in Figure 1, the first involving a specified functional dependence and the other based simply on probability mass assignments made at the extreme values of feature intensity and estimated through linear interpolation elsewhere. The purpose of these probability mass functions was to convert feature intensity values on a point-by-point basis to probability assignments for each of the focal elements of the frame of discernment for every point in

the survey grid at which the feature layer was defined. At points on the grid where feature intensity values were missing, the algorithm was programmed to assign all probability mass to uncertainty, reflecting that this situation represents a lack of knowledge. This process results in a site-wide map of probability mass assignments for each focal element for each feature layer.

These resulting probability mass assignments were then combined on a point-by-point basis throughout the survey grid according to Dempster's rule of combination, as illustrated in Figure 2:

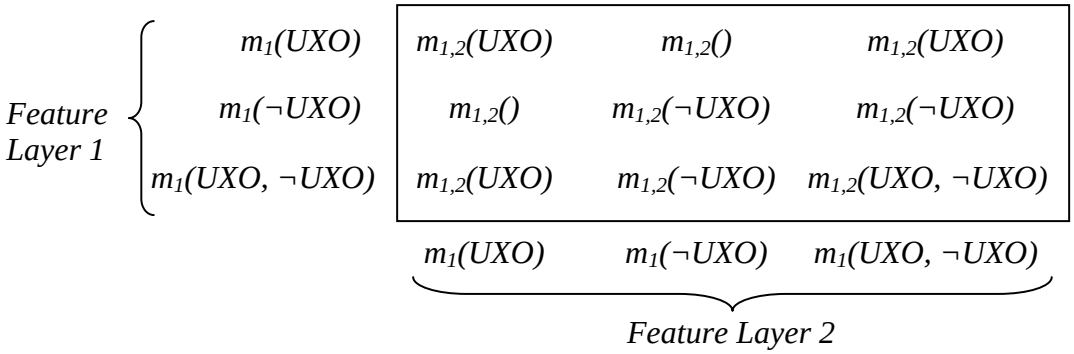



Figure 2. Dempster's rule of combination implemented for a UXO assessment application.

At each point in the survey map, the cross products of the probability mass assignments from the three focal elements of the UXO frame of discernment form a set of nine terms. Three of these support the focal element {*UXO*}, three support {¬*UXO*}, one supports {*UXO*, ¬*UXO*}, and two represent evidence assigned to null set, {}, which indicates the amount of conflict between the different lines of evidence. The output probability masses for {*UXO*}, {¬*UXO*}, and {} were calculated by summing the appropriate terms. Evidence was combined in a serial fashion, with the probability mass assignments of additional layers combined with the output assignments generated by Dempster-Shafer combination of the previous layers. In order to preserve the transitive nature of this operation (i.e., to ensure that evidence combined in any order provided the same output) assignments made to the null set were propagated separately until the final feature layer was combined. Final output was generated by unit normalization of the non-empty frame element assignments: {*UXO*}, {¬*UXO*}, and {*UXO*,¬*UXO*}.

With the final output, decisions regarding area delineation can be made that take into account both the weight of evidence attributed to UXO being present and the weight assigned to uncertainty. The values of the Dempster-Shafer quantities of support and plausibility for focal element {*UXO*} at each point in the survey map can be calculated as, respectively, the output assignment to {*UXO*} itself, and one minus the output assignment to {¬*UXO*}, or, equivalently, the sum of the assignments output to {*UXO*} and {*UXO*,¬*UXO*}.

# Results and Discussion

Work in year two centered on data fusion algorithm and architecture development. Concepts in feature layer development begun in year one were extended and tested against new data sets acquired from wide-area assessment of the former Kirtland Bombing Targets N1 and N3. The performance of an NRL-developed, crater-detection algorithm on the Pueblo and Kirtland sites was examined as well as compared to manually-located craters generated by the survey teams. Various algorithmic approaches for fusing feature layers into UXO assessments were examined and a Dempster-Shafer formalism was chosen for eventual implementation and evaluation of a software prototype. Finally, approaches for implementing specific heuristic rules to incorporate particular feature interdependencies were developed and implemented in the prototype data fusion framework.

**Generation of feature layers**

Feature layer developments in year two of project MM-1510 focused on extending techniques developed in year one to data from new sites, and in consolidating a set of available features for each site that could be used for further data fusion algorithm development. Early in the year, data acquired during the ESTCP WAAPP survey of former Kirtland Bombing Targets N1 and N3 were obtained through SERDP. Roughly commensurate with the data acquired from the Pueblo precision bombing range that were obtained in year one of project MM-1510, the Kirtland data streams were imported in a similar fashion by first generating a common map grid of the survey area at one meter resolution and then registering each data set to it. [11]

Figure 3 depicts the surface generated via interpolation of helicopter magnetometry data acquired from the Kirtland ESTCP WAAPP site. As can be seen, data at the Kirtland site were acquired in two separate continuous regions, a northern section and a smaller southern section.

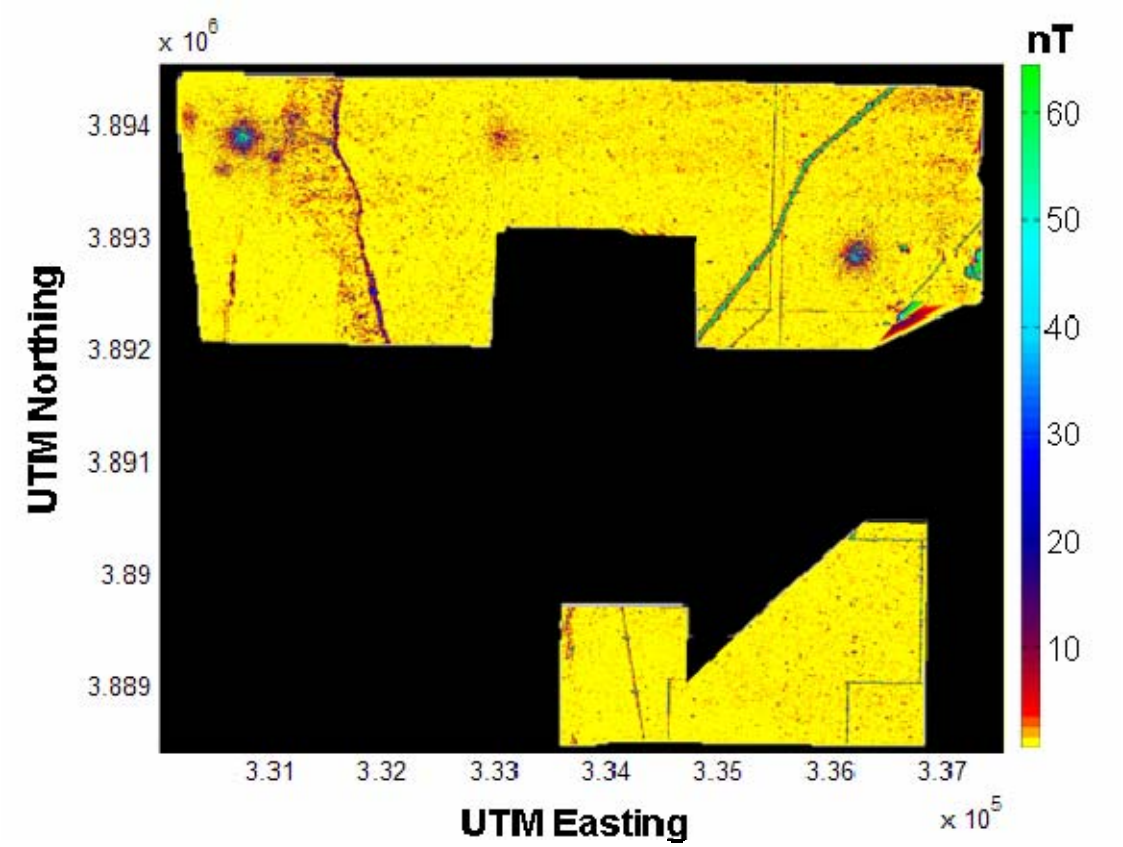



Figure 3. Helicopter magnetometry survey from the Kirtland ESTCP WAAPP survey site.

Figure 4 depicts the corresponding aerial LiDAR survey while Figure 5 depicts the aerial orthophotographic survey acquired at the same time.

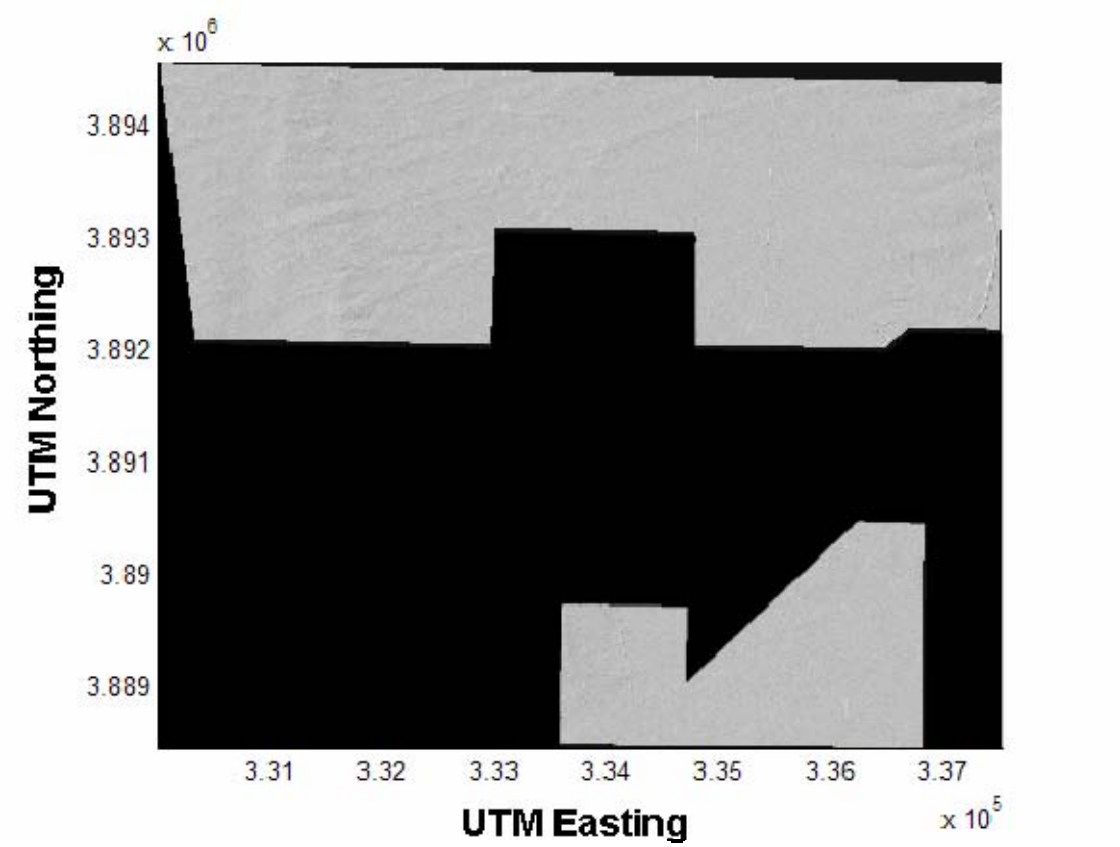



Figure 4. LiDAR survey from the Kirtland ESTCP WAAPP survey site.

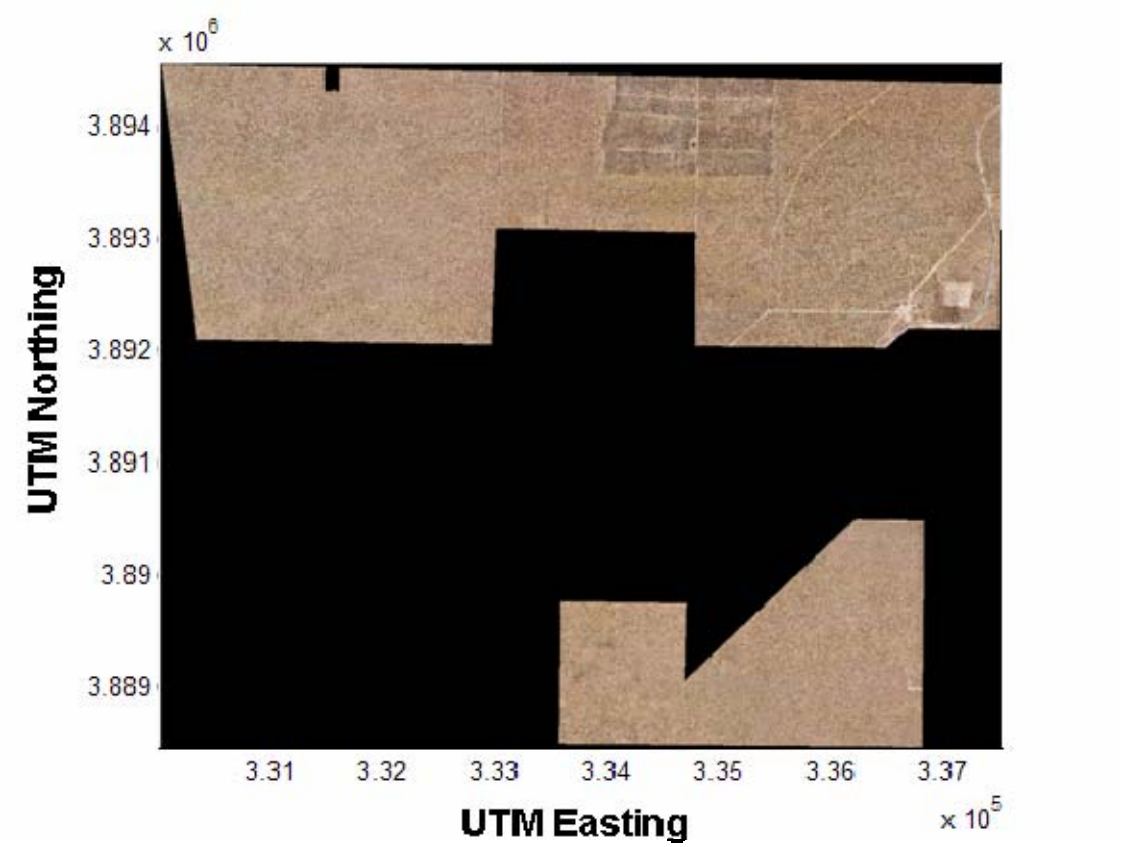



Figure 5. Aerial orthophotographic survey from the Kirtland ESTCP WAAPP survey site.

Following importation, feature extraction algorithms developed in year one for the Pueblo site data were then examined against the newly imported Kirtland data. A comparison was made of auto crater detection algorithm performance between the two sites. A set of manually identified crater locations was provided for Kirtland as part of the data acquired through SERDP. A comparison of manually located craters versus craters located automatically via the crater detection algorithm is shown in Figure 6. Figure 7 depicts feature intensity maps generated from these two feature sets. For comparison, a map of automatically identified crater locates and the resulting feature intensity maps are shown in Figure 8.

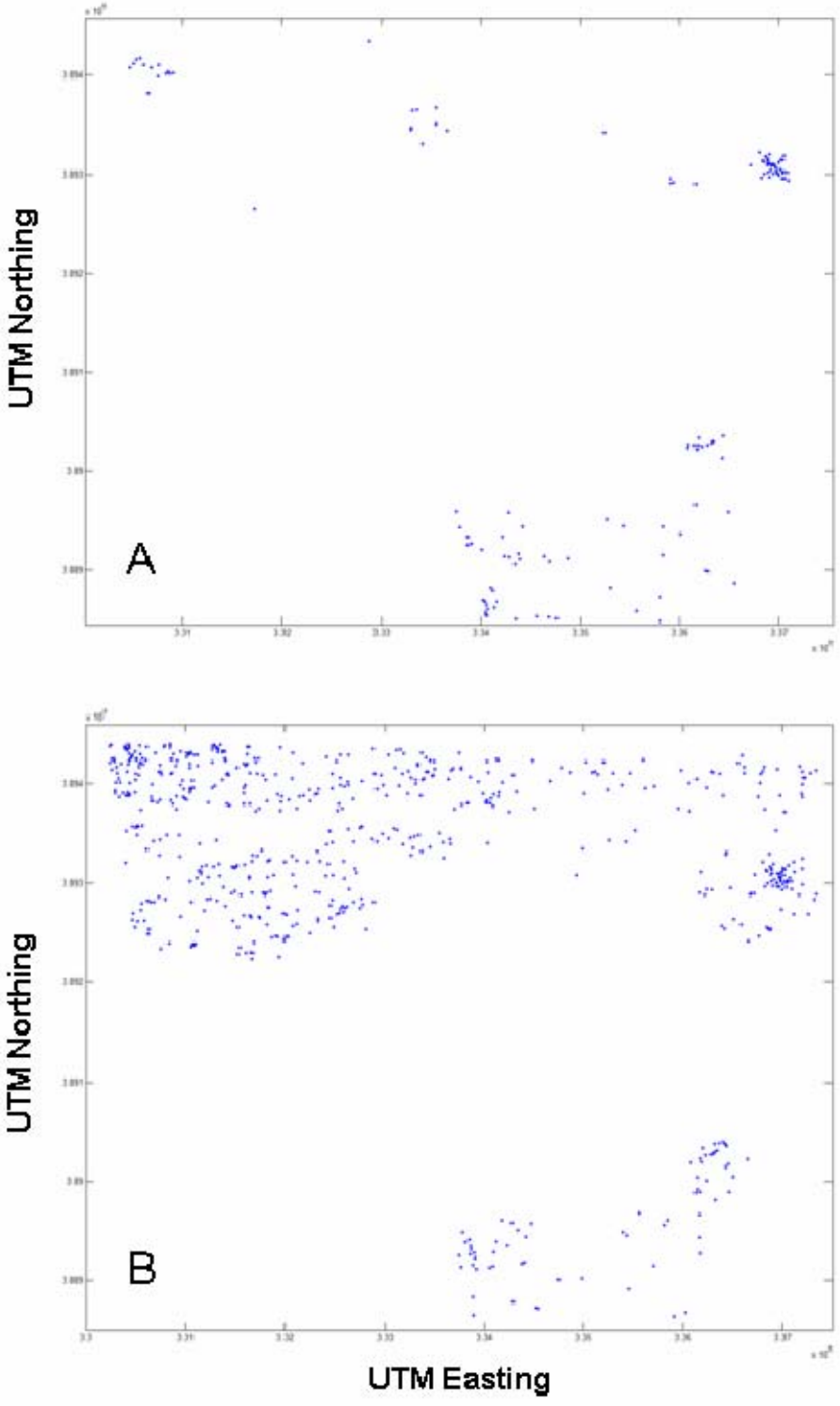



Figure 6. Craters detection at the Kirtland site. (A) depicts manually identified craters (Versar, Inc.) (B) depicts automatically identified craters via an NRL crater-detection algorithm.

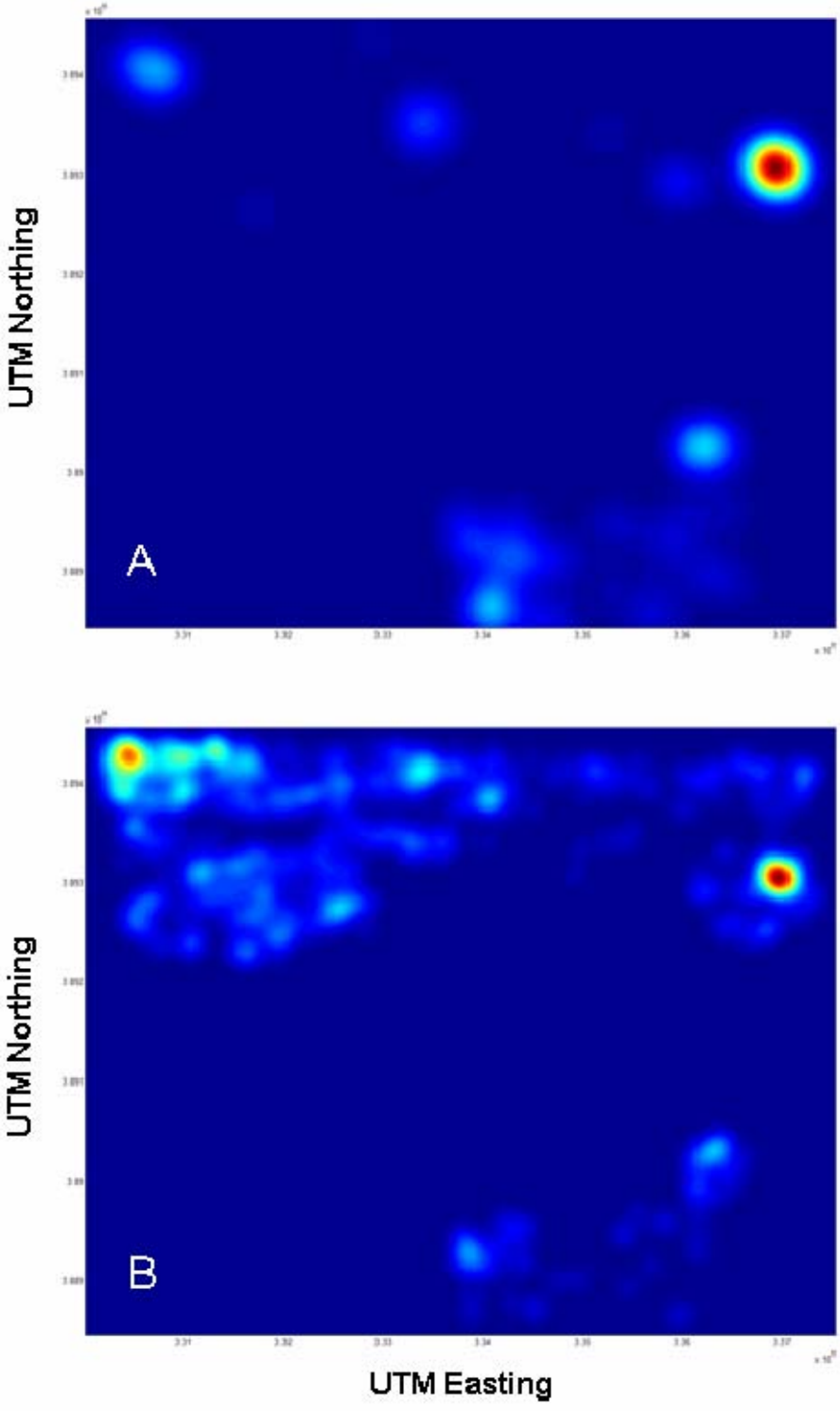



Figure 7. Crater density estimate determined via Kernel Denisty Estimation. (A) depicts manually identified craters (Versar, Inc.). (B) depicts automatically identified craters via an NRL crater-detection algorithm. Colors are scaled from blue (=0) to red (=1).

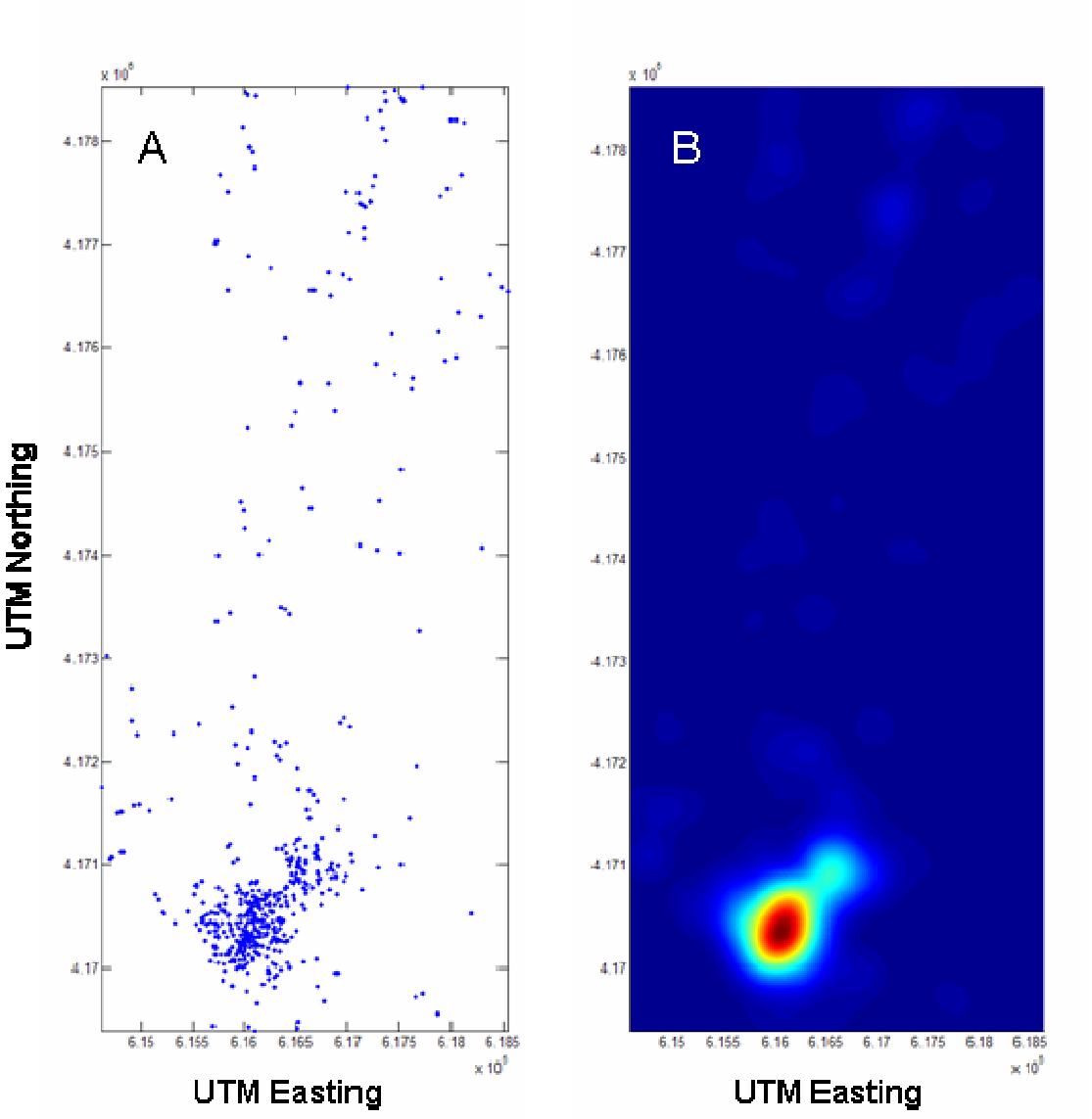



Figure 8. Crater detection at the Pueblo site. (A) depicts automatically identified craters via an NRL crater-detection algorithm. (B) depicts a crater density estimate determined via Kernel Density Estimation. Colors are scaled from blue (=0) to red (=1).

The results of this comparison indicated that algorithm thresholds for maximum detection rates were similar from site to site, although the surface topography presented by the Kirtland site made accurate crater detection more challenging as it presented greater background interference. Thus, the chief difference between the results attained at the two sites was that the Kirtland site demonstrated higher apparent false positive rates, and, unlike Pueblo, these false positive rates did not correlate well with algorithm threshold setting. At either site, the addition of a wavelet-based filter enhanced automatic crater detection, although in the absence of reliable truth data, the improvement and false positive rates were difficult to quantify.

In addition to crater locations at the Kirtland site, various manually delineated regions of interest were provided for both the Pueblo and Kirtland sites. These were generated by ESTCP WAAPP performers, or aggregated by those performers from other third-party sources. The individual regions were provided as ESRII "shape files" containing UTM geocoordinates for the boundaries described by the region and, typically, a manually entered text label describing the nature of the region. These labels included both highly specific terms and rather vague characterizations. In order to generate appropriate feature layers, regions were grouped according to their relationship to UXO. Thus, all regions describing visible bombing targets were combined into one feature layer, all regions characterized as a "potential MRA" (munitions remediation area) were grouped into a second layer, and all those describing features that could potentially lead to significant non-UXO related magnetometry signal (e.g. "structure", "fence line", "road") were grouped into a third layer. Other regions with more vague labels (e.g. "linear feature"), more uncertain relationships to either the presence or absence of UXO (e.g. "bumps", "berms", and "depressions") were grouped into other feature layers, but were not used for subsequent data fusion in this work. Figures 9 and 10 depicts feature layers generated for the Kirtland site.

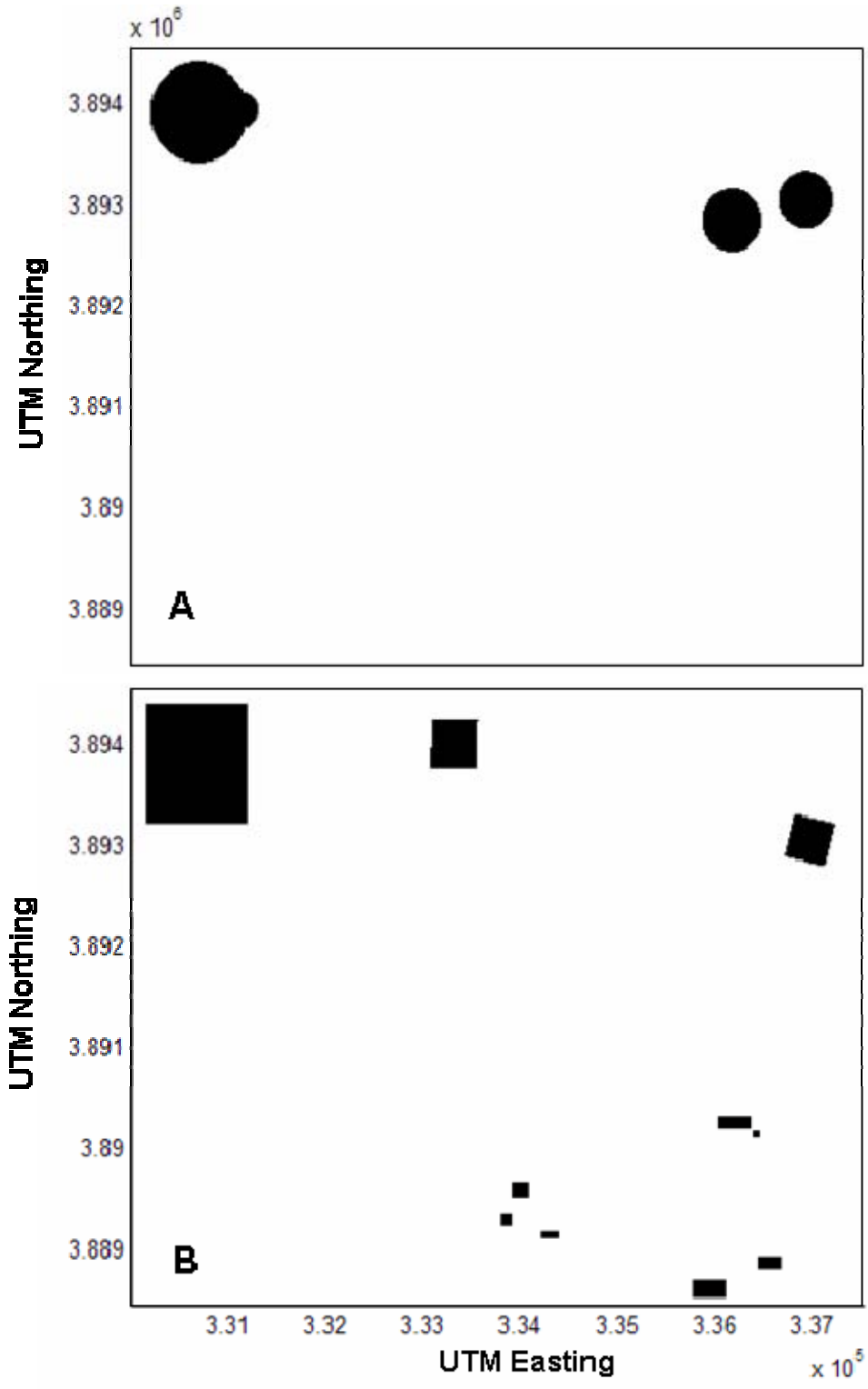



Figure 9. Feature layers resulting from (A), manually identified ship targets dilated to a larger effective area, and (B), manually identified munitions areas.

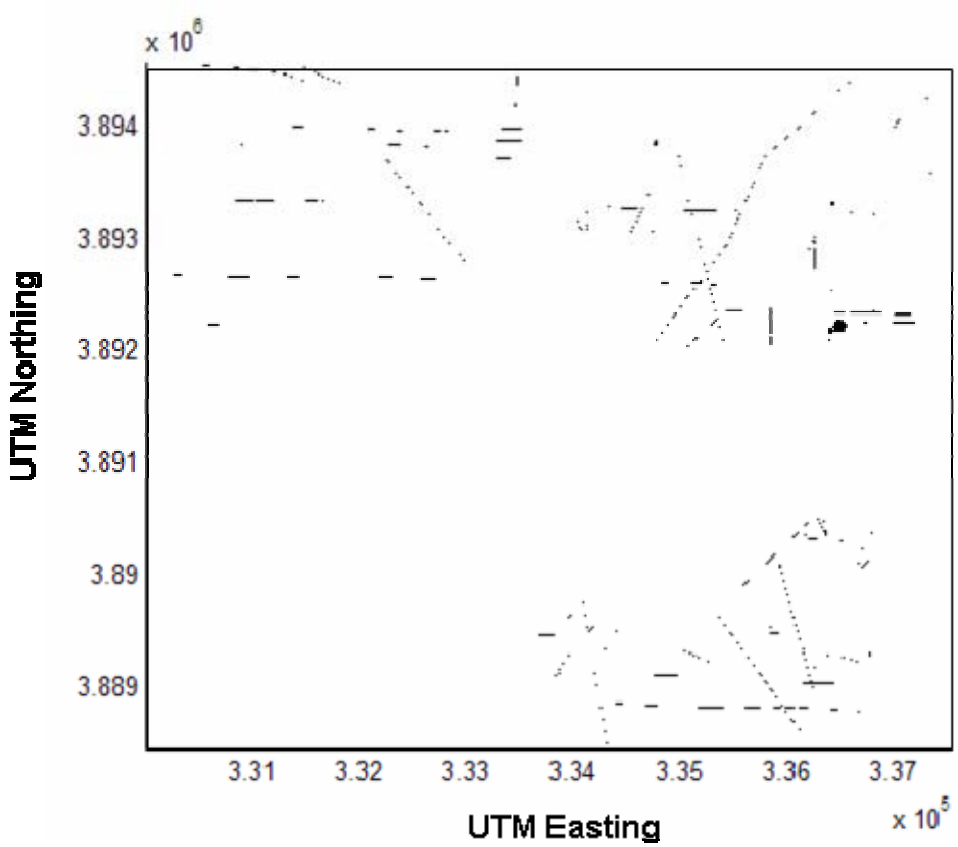



Figure 10. Feature layers resulting from manually identified roads, fences, pipelines, and man-made structures within the Kirtland site.

An enlarged region in Figure 11 demonstrates the process of region dilation to account for a given feature's area of influence. Visible bombing targets were expanded to include a circle of a diameter twice the length of the major axis of, and centered on, the original target. A similar process is necessary for any manually identified region of interest in which the area of influence of that feature exceeds the boundaries of the region itself.

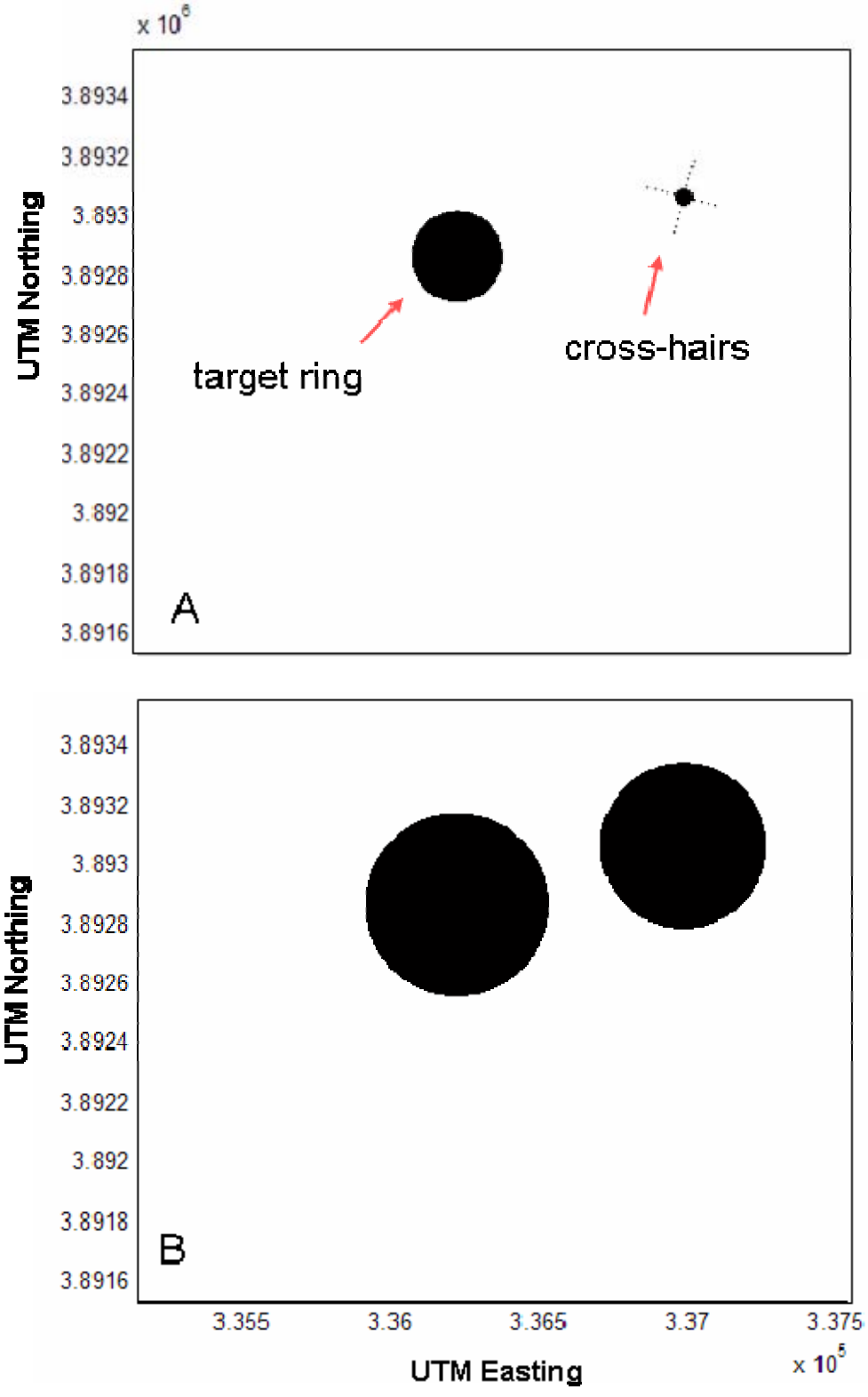



Figure 11. Enlarged region of the Kirtland feature layer containing bombing targets (A) before, and (B), after dilation.

A similar procedure utilizing manually delineated regions of interest at the Pueblo site resulted in three feature layers: manually identified ship targets, manually identified munitions areas, and manually identified man-made structures. The layers are depicted in Figure 12.

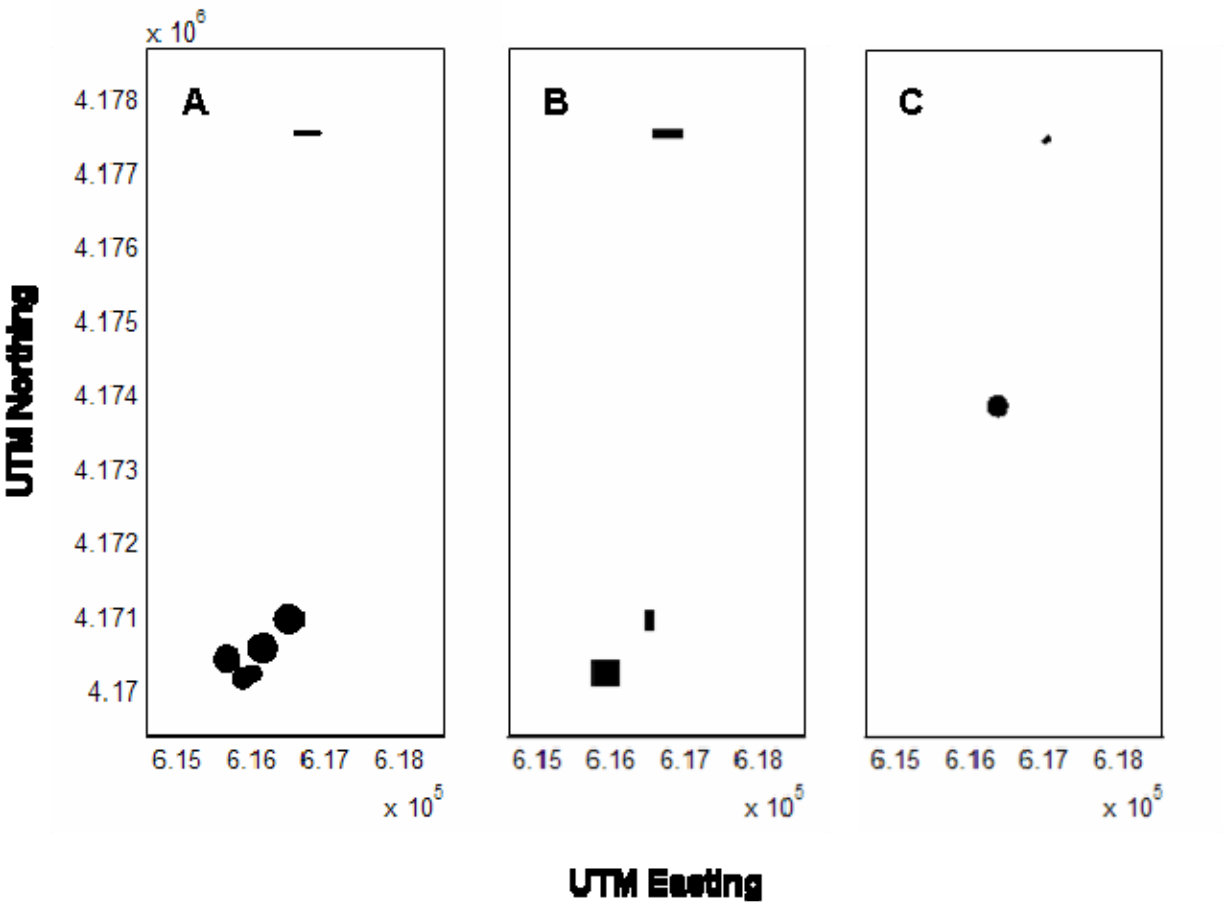



Figure 12. Feature layers resulting from (A), manually identified ship targets, and (B), manually identified munitions areas, and (C), manually identified man-made structures in the Pueblo site.

So called "magnetic anomalies" were algorithmically generated from helicopter magnetometry data by Sky Research, Inc., the ESTCP WAAPP performer that acquired helicopter magnetometry data at both the Pueblo and Kirtland sites. Each anomaly indicates a survey grid location at which the acquired magnetometry signal was consistent with the presence of a UXO-related object. These anomaly locations were obtained from SERDP and were utilized to create corresponding feature layers consisting of anomaly density estimate maps for both the Pueblo and Kirtland sites. These maps were generated using the same techniques employed in generating crater density maps from detected crater locations.

Sky Research supplied a list of anomaly locations identified algorithmically from the helicopter magnetometry survey data. From a ROC curve analysis, Sky Research selected a threshold that resulted in a "false positive rate" of approximately 25% when

compared to manually identified anomalies at the Pueblo and Kirtland sites. [21] The automatically identified anomalies are shown in Figure 13A for Kirtland and Figure 14A for Pueblo. Both figures exhibit a nearly uniform spatial distribution of detected anomalies punctuated by clusters of anomalies near known targets and gaps at known locations of roads and pipelines. A histogram of the minimum (nearest neighbor) distances between magnetic anomalies at the Kirtland site is shown in Figure 15. The data were distributed similar to a Poisson distribution with a mean and variance of 6.0 with additional data points present in the tail region (i.e., larger distances between anomalies.)

A nearest neighbor clustering algorithm was developed to distinguish anomalies associated with UXO from those due to background. The algorithm was built on the assumption that anomalies due to non-UXO phenomena were likely to be more uniformly distributed throughout the surveyed areas, due to their random origin, than UXO-related anomalies, which were more likely to be clustered around bombing targets or areas replete with magnetic rocks. Application of the automatic clustering algorithm to filter helimag anomaly feature sets demonstrated significant reduction in background signal, as shown in Figures 13 and 14, and in the associated feature density maps, shown in Figures 16 and 17. The algorithm removed anomalies with 10 or fewer neighboring anomalies within a circle of radius 64 meters. Anomalies passing these criteria were highly clustered around known target areas, though some clusters were known to be due to the presence of magnetic geology.

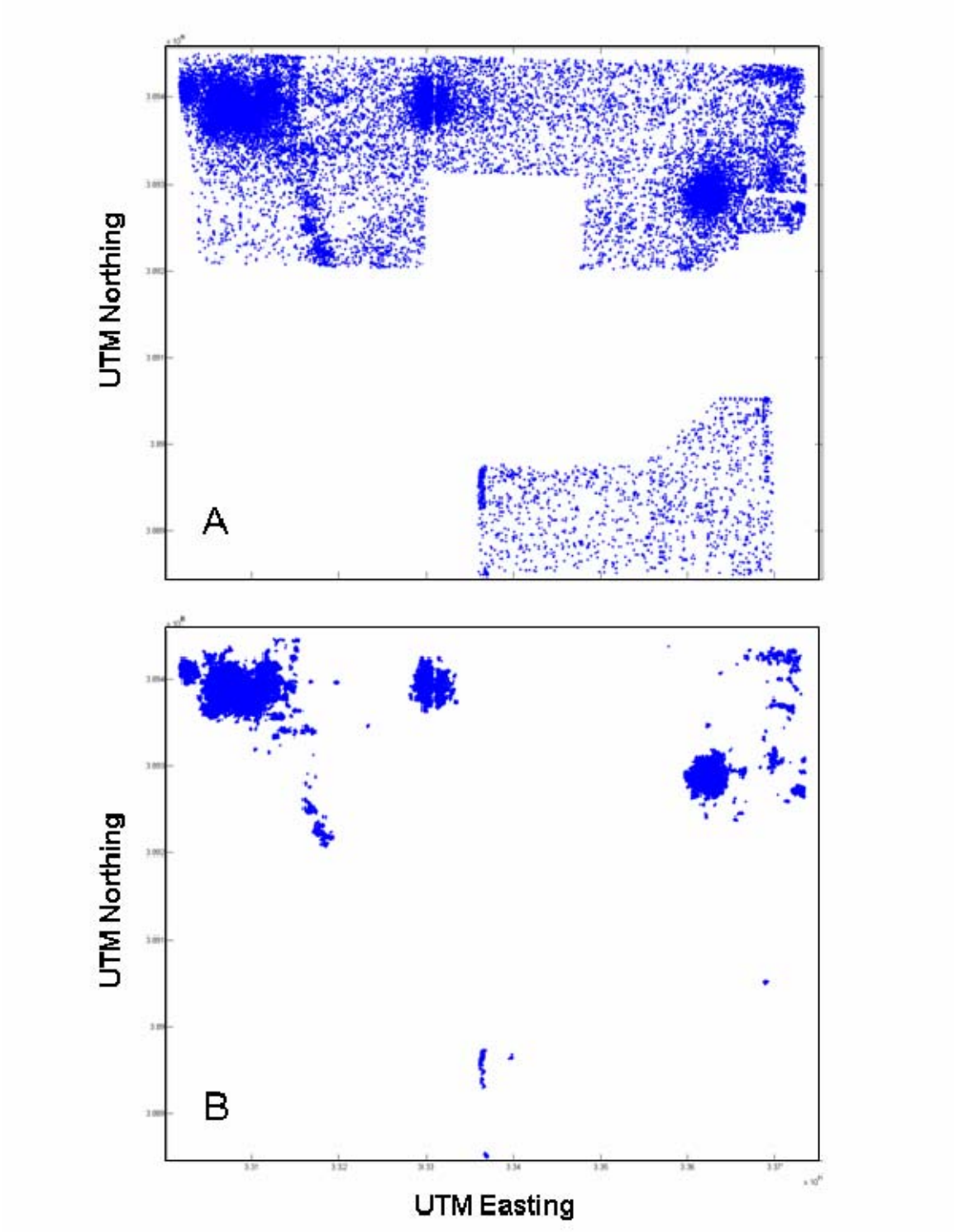



Figure 13. Magnetic anomalies at the Kirtland site. (A) depicts algorithmically identified anomaly locations (Sky Research). (B) depicts anomaly locations after filtering with an NRL clustering algorithm.

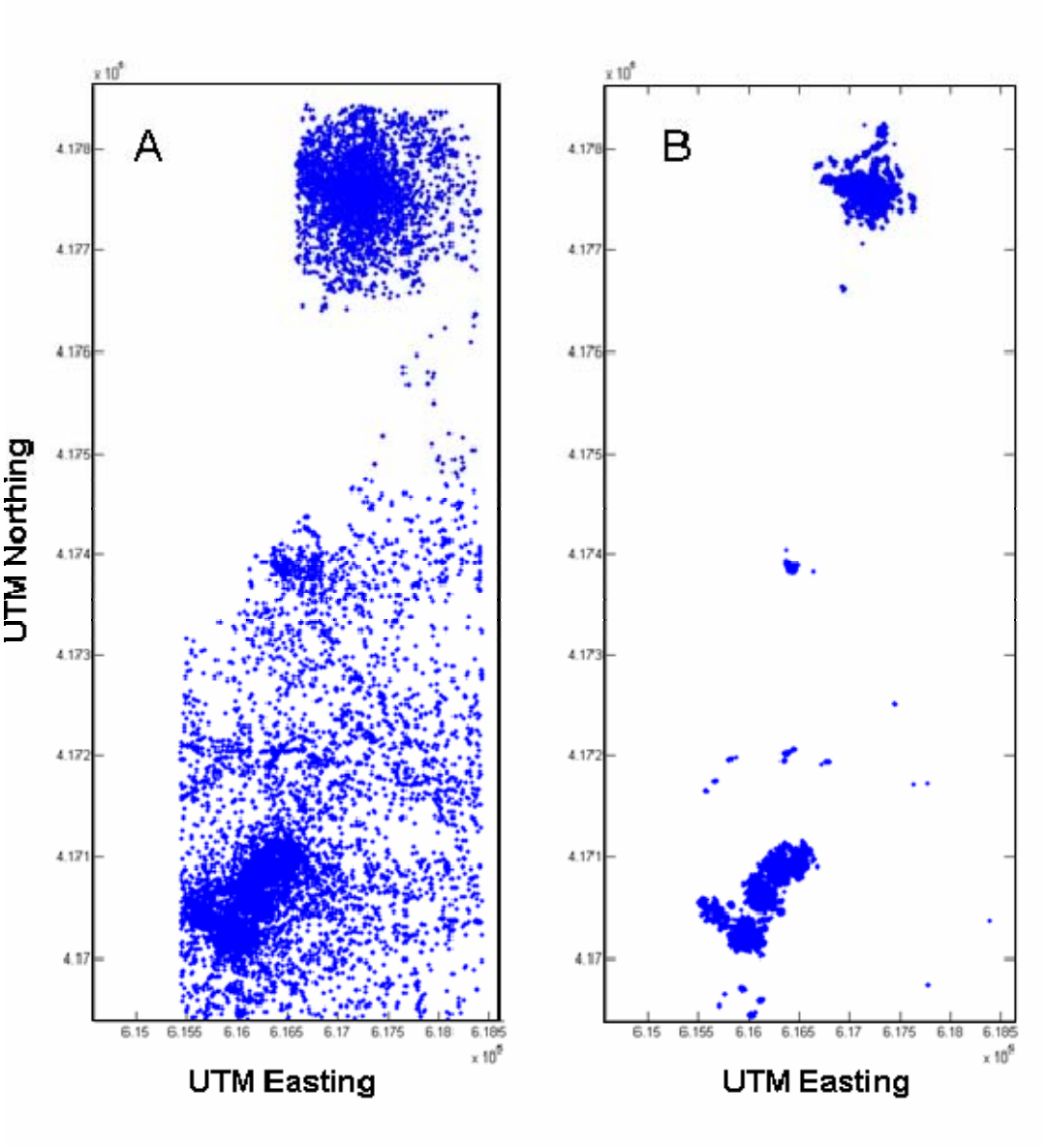



Figure 14. Magnetic anomalies at the Pueblo site. (A) depicts algorithmically identified anomaly locations (Sky Research). (B) depicts anomaly locations after filtering with an NRL clustering algorithm.

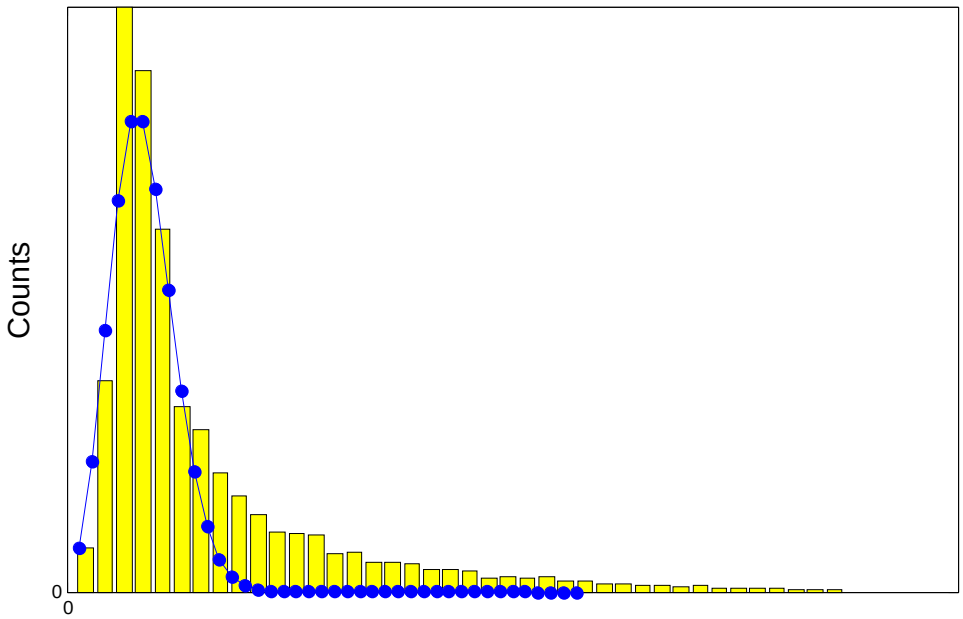



Figure 15. Histogram of nearest neighbor distances between magnetic anomalies at the Kirtland site (yellow bars), overlay of Poisson distribution of mean 6.0 (= variance) (blue line).

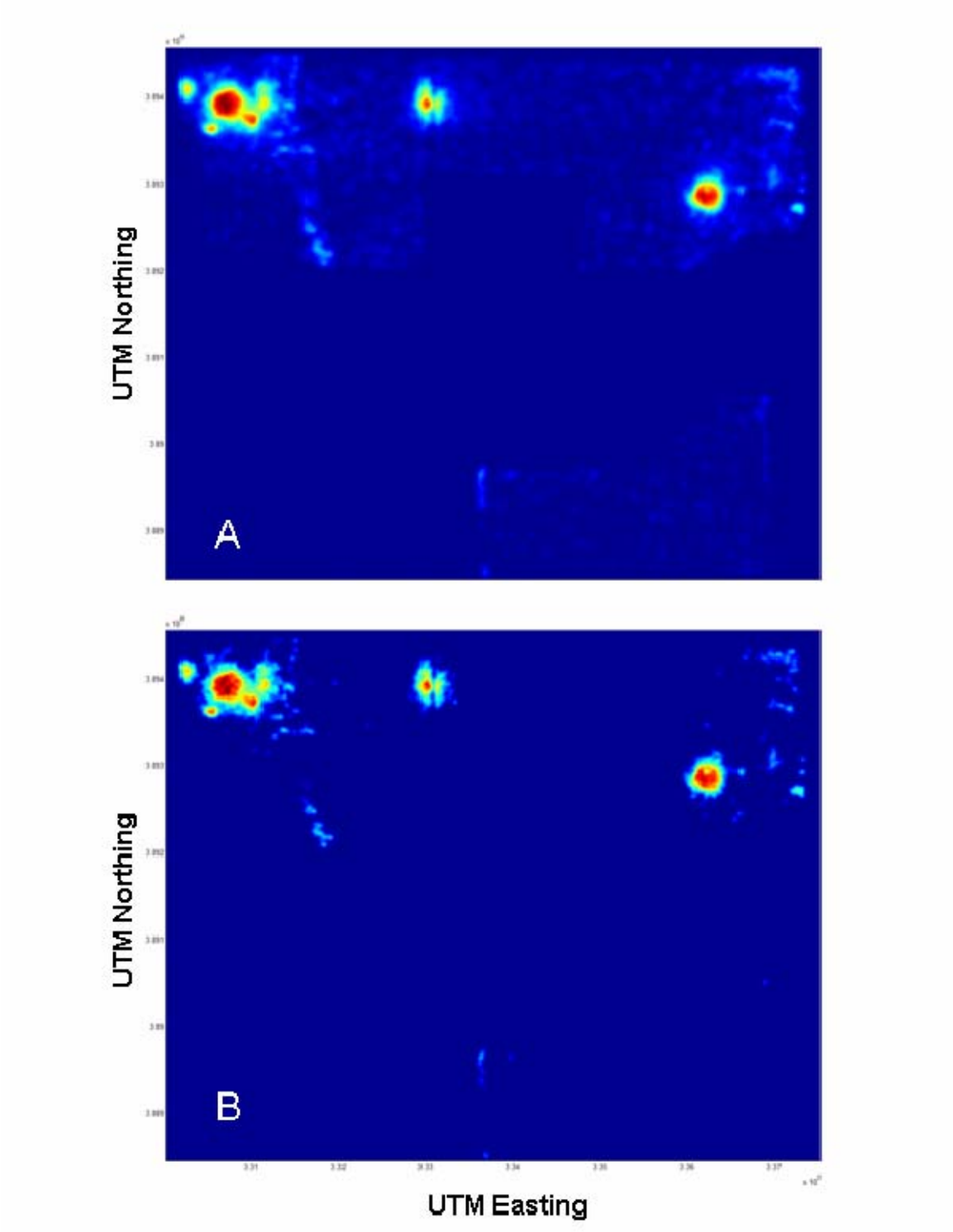



Figure 16. Magnetic anomalies at the Kirtland site. (A) depicts estimated anomaly density for Fig. 13A. (B) depicts estimated anomaly density for Fig. 13B.

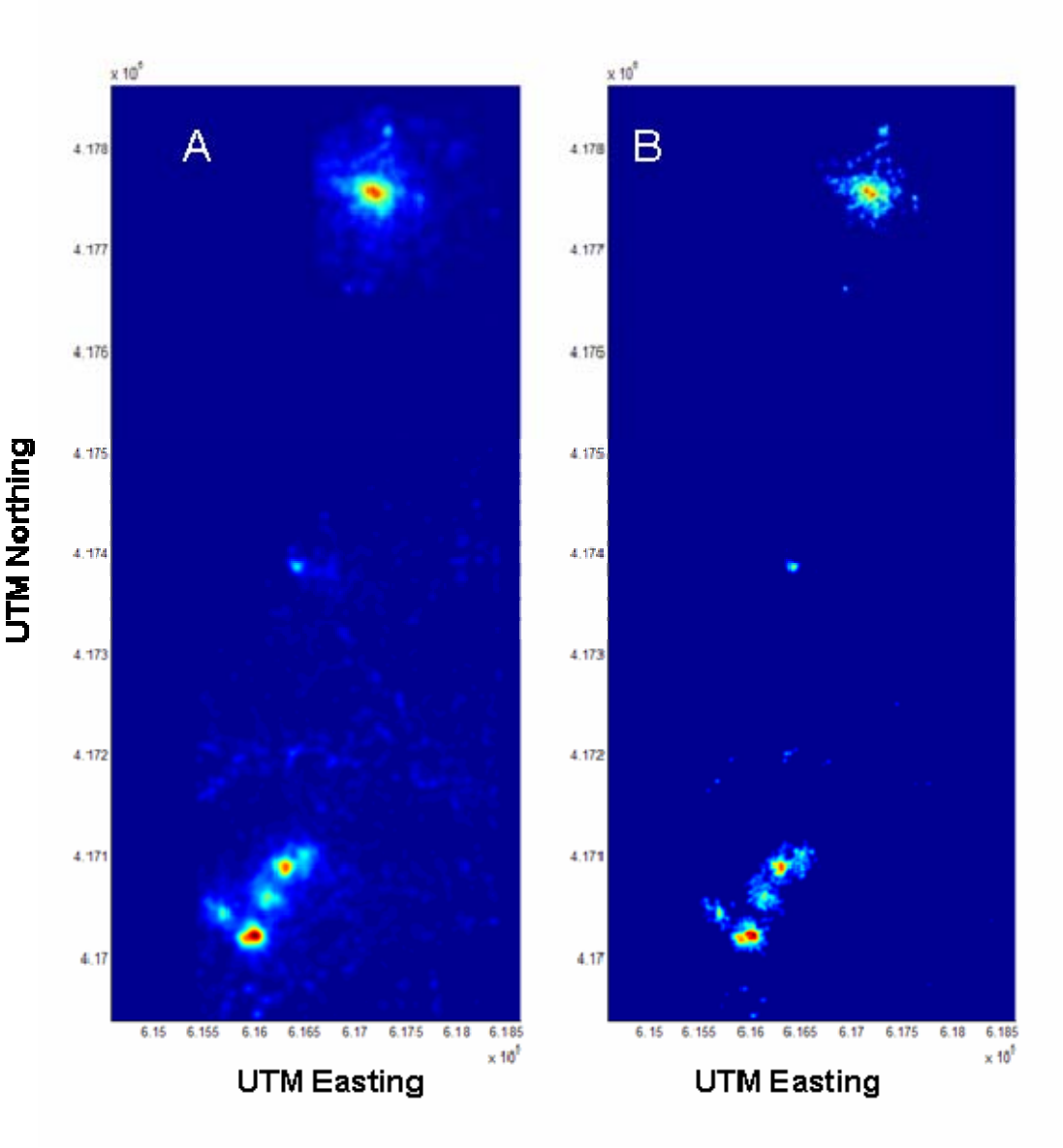



Figure 17. Magnetic anomalies at the Pueblo site. (A) depicts estimated anomaly density for Fig. 14A. (B) depicts estimated anomaly density for Fig. 14B.

Following feature layer development and generation, the following feature data layers were available for data fusion:

- Density of helicopter magnetometry anomalies acquired from Sky Research, both with and without NRL nearest-neighbor clustering filter.
- Density of threshold-applied, morphologically filtered magnetometry signals
- Density of automatically detected craters
- Density of manually detected craters from Versar (Kirtland only)
- Manually delineated bombing targets from Versar
- Manually delineated munitions areas from Versar

- Manually delineated man-made structures from Versar

In general, the feature layers generated fell into one of two types: First, layers that were direct indications of the presence of UXO, such as the magnetometry data and anomalies for which $p$(UXO) and $p(\neg$UXO) were well-defined. And second, those layers that were indirect indications of the presence of UXO, for which only $p$(UXO) was well-defined. These latter feature layers included crater anomalies, manually identified target areas, munitions areas, and man-made structures.

**Data fusion results**

*Heuristic Approach.* The output of heuristic-based fusion of a subset of available data layers is shown in Figure 18. In this example, the output was the result of a linear combination of intensity values from the magnetic anomaly density, manual crater density, bombing target, and munitions area feature layers available for the Kirtland site. In essence, the linear combination of layers is functionally equivalent to the overlay of multiple feature layers and represents roughly the same level of knowledge that simple visualization tools provide. The output denotes areas of likely UXO contamination, but says nothing about the likelihood of UXO contamination in those areas.

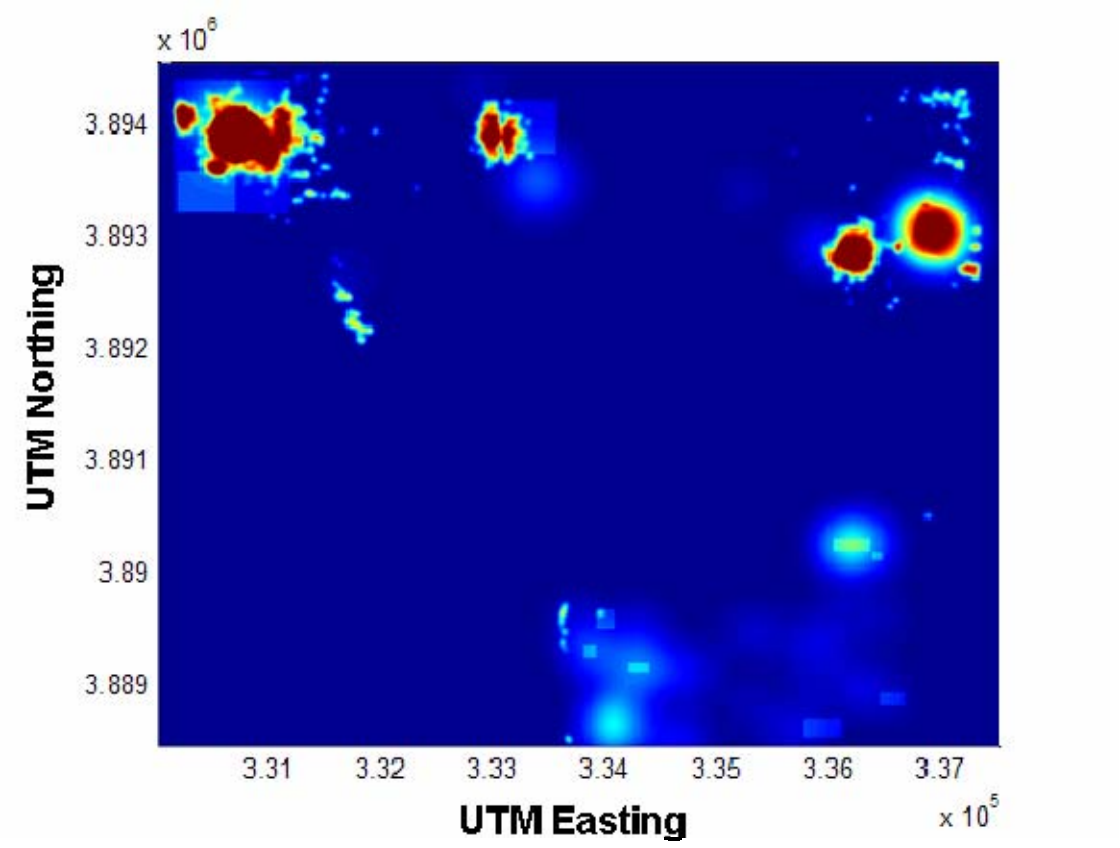



Figure 18. Heuristic-based data fusion output utilizing magnetic anomaly density, manually detected crater density, bombing target, and munitions area feature layers. (Map is color scaled from blue=0 to red=1).

*Bayesian Approach.* A Bayesian algorithm implementation was tested utilizing similar subsets of feature layers: manually delineated bombing targets and munitions areas, automatically detected crater density, and magnetic anomaly density for the Pueblo site, and manually delineated bombing targets and munitions areas, manually detected crater density, and magnetic anomaly density for the Kirtland site. Each feature layer required the specification of conditional probabilities as a function of intensity value. Initial assignments to the conditional probability distributions were made, reflecting a subjective assessment of the strength of association between each feature layer and the presence of UXO. Assignments were made to the extremal values of feature intensity (0 and 1), and intermediate values were calculated through linear interpolation.

| | Feature Intensity | p(f\|UXO) | p(f\|¬UXO) |
|---|---|---|---|
| Crater Density | 0 | 0.1 | 0.1 |
| | 1 | 0.7 | 0.2 |
| Helimag Anomaly Density | 0 | 0.01 | 0.8 |
| | 1 | 0.7 | 0.01 |
| Target Features | 0 | 0.5 | 0.5 |
| | 1 | 0.8 | 0.2 |
| Munitions Areas | 0 | 0.5 | 0.5 |
| | 1 | 0.7 | 0.3 |

Figure 19. Initial conditional probability assignments for Bayesian data fusion combining helimag anomaly density, crater density, known bombing targets, and manually delineated munitions areas.

The prior probability of UXO, $p$(UXO), was set to 0.5, and reflected complete ignorance as to the presence or absence of UXO at the site before the inclusion of any observational evidence. This prior assignment was updated serially with each new feature layer, as described in the Method section. In this process, the specified conditional probability distribution was used to convert the feature intensity values of each layer into corresponding values of $p(f \mid UXO)$ and $p(f \mid \neg UXO)$, and on a point-by-point basis, a posterior probability of UXO, $p(UXO \mid f_1, f_2)$ was calculated using the prior and the two conditional probabilities. The resulting output maps are shown in Figures 20 and 21.

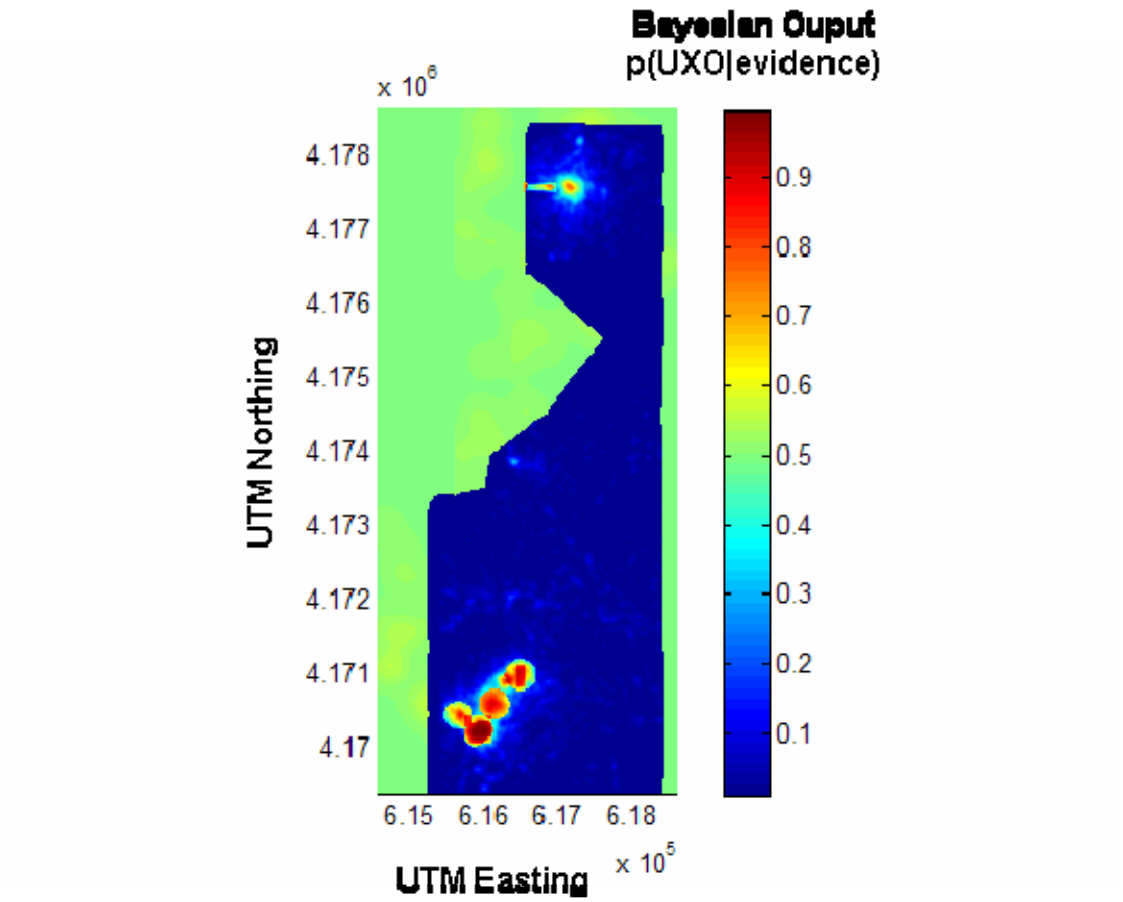



Figure 20. Bayesian data fusion output for the Pueblo site utilizing magnetic anomaly density, automatically detected crater density, bombing target, and munitions area feature layers.

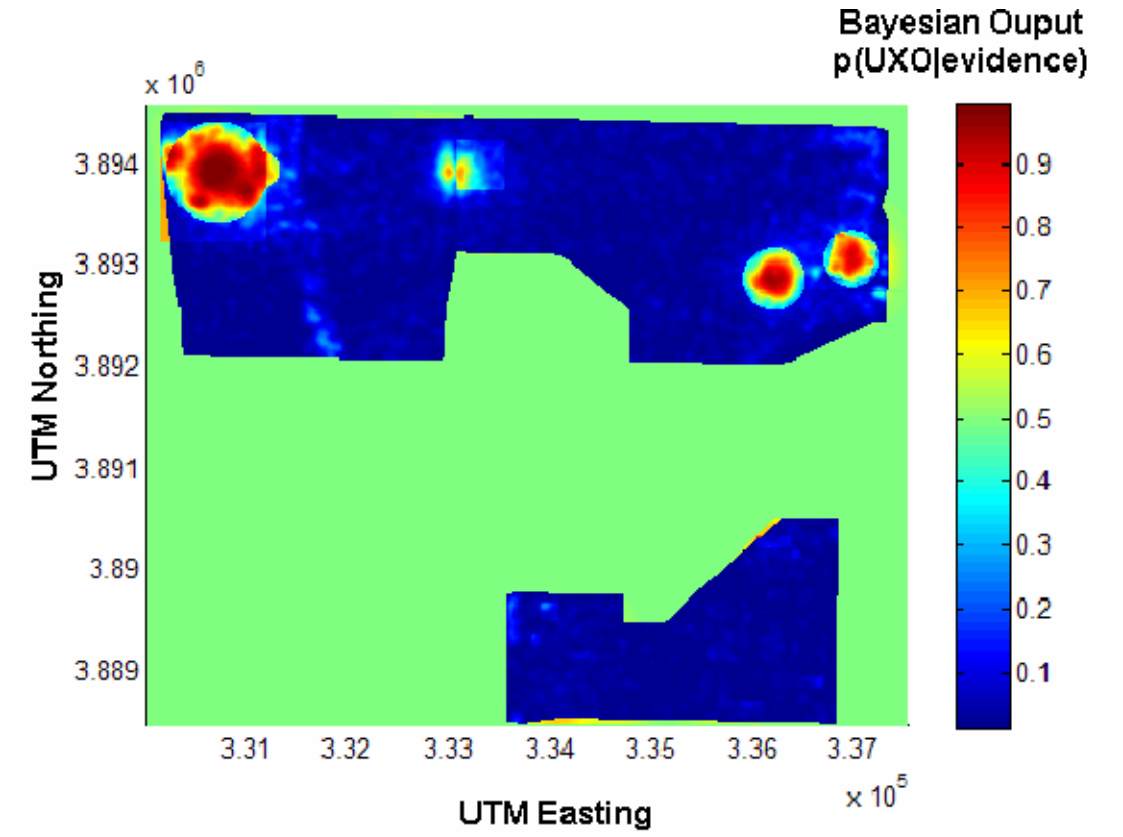



Figure 21. Bayesian data fusion output for the Kirtland site utilizing magnetic anomaly density, manually detected crater density, bombing target, and munitions area feature layers.

While the output appears reasonable, the conditional probability assignments are problematic for two reasons. First, conditional probability assignments for the UXO application are difficult to estimate in an intuitive fashion based on subjective knowledge. Second, meaningful objective conditional probability distributions are impractical to obtain through empirical means.

The first problem arises when a feature layer has an indirect or poorly defined relationship to UXO. As an example, consider the implications of the probability assignments made to the bombing target feature layer. Subjectively, one supposes that grid points within bombing target features are more predisposed to contamination by UXO or UXO-related scrap than grid points outside of target features. Assuming that this is the extent of our knowledge regarding the relationship between bombing targets and UXO, it is not unreasonable to expect that, in the absence of any other information, grid points outside of target features are as likely to be contaminated with UXO as not. That is, the fact that a grid point is *not* located within a bombing target feature does not convey any information regarding UXO contamination.

Unfortunately, it is difficult to render the subjective knowledge presented by the bombing target feature layer into assignments for conditional probabilities. Grid points within the feature layer are restricted to one of two values: 1 or 0. The former indicates a pixel is

within a bombing target region, the later indicates that it is not. The corresponding conditional probability functions thus have only two assignments, one for each possible feature layer value. For $p(f \mid UXO)$, these values should correspond to the proportion of the total number of UXO-contaminated grid points at the site that are within the bombing targets and those that are outside of bombing targets, respectively. For $p(f \mid \neg UXO)$, these values should correspond to the proportion of non-contaminated grid points that are within and without bombing targets, respectively. To make conditional probability assignments in a principled fashion, a great deal more information is required regarding the nature of the relationship between the feature and UXO. Attempting to estimate values that force behavior agreeing with subjective knowledge results in nonsensical assignments. In order to bring about the proper behavior for grid points outside of target areas, $p(f \mid UXO)$ must equal $p(f \mid \neg UXO)$ where the feature layer equals zero. Unfortunately, this means that $p(f \mid UXO)$ must also equal $p(f \mid \neg UXO)$ where the feature layer is one, as the conditional probability distributions must integrate to unity over all possible feature values. Assignment of conditional probability values that represent a predisposition of UXO contamination for grid points within bombing targets, which reflects our subjective knowledge of the feature, thus forces either the abandonment of probabilistic tractability (i.e., conditional probability distributions that don't integrate to unity,) or of the desired treatment of grid points outside of bombing targets (i.e., that such information represents complete ignorance as to the presence or absence of UXO.)

For feature layers with a more direct relationship to UXO, it is more likely that meaningful conditional probability functions can be estimated through empirical evidence. For example, at the Pueblo site, a limited amount of ground truth data were obtained by SERDP in which UXO detections via surface magnetometry were physically examined by digging at each detect site. These surveys included regions suspected to contain UXO, as well as one region suspected to be free of UXO. The resulting 621 detects were characterized as UXO or UXO-related scrap, non-UXO scrap, or geologic (i.e., no object was located during the dig.) Extracting the magnetic anomaly density at each dig location yields the histograms shown in Figure 22.

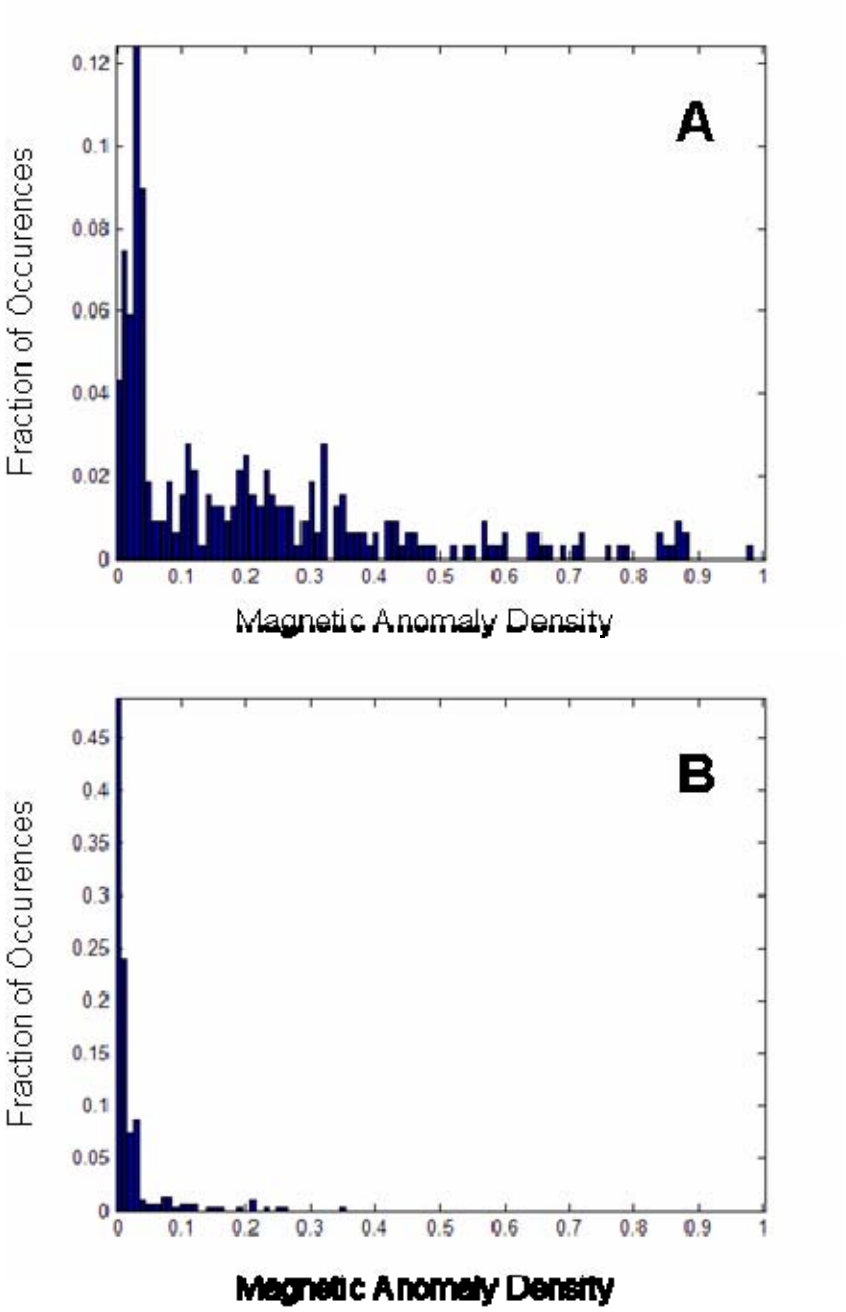



Figure 22. Relationship between ground truth survey data from the Pueblo site and magnetic anomaly density feature layer. A) depicts the distribution of density values for locations where UXO or UXO-related objects were found while B) depicts the distribution of density values for locations determined to be free of UXO.

Such data could be used to generate estimates of $p(f \mid UXO)$ and $p(f \mid \neg UXO)$, but these implicitly assume that the sampling employed in the ground truth survey represented an unbiased sampling of the contaminated and non-contaminated regions of the site. Such an assumption is likely to prove false, due to the limited scope of the survey and due to

the fact that only regions resulting in a "detect" from surface magnetometry were dug up, resulting in a bias of the data towards UXO-contaminated regions. A truly randomized sampling of the site with the intent to produce more accurate estimates of $p(f \mid UXO)$ and $p(f \mid \neg UXO)$ is likely to be prohibitively expensive in terms of time and cost, negating the utility of an empirical approach.

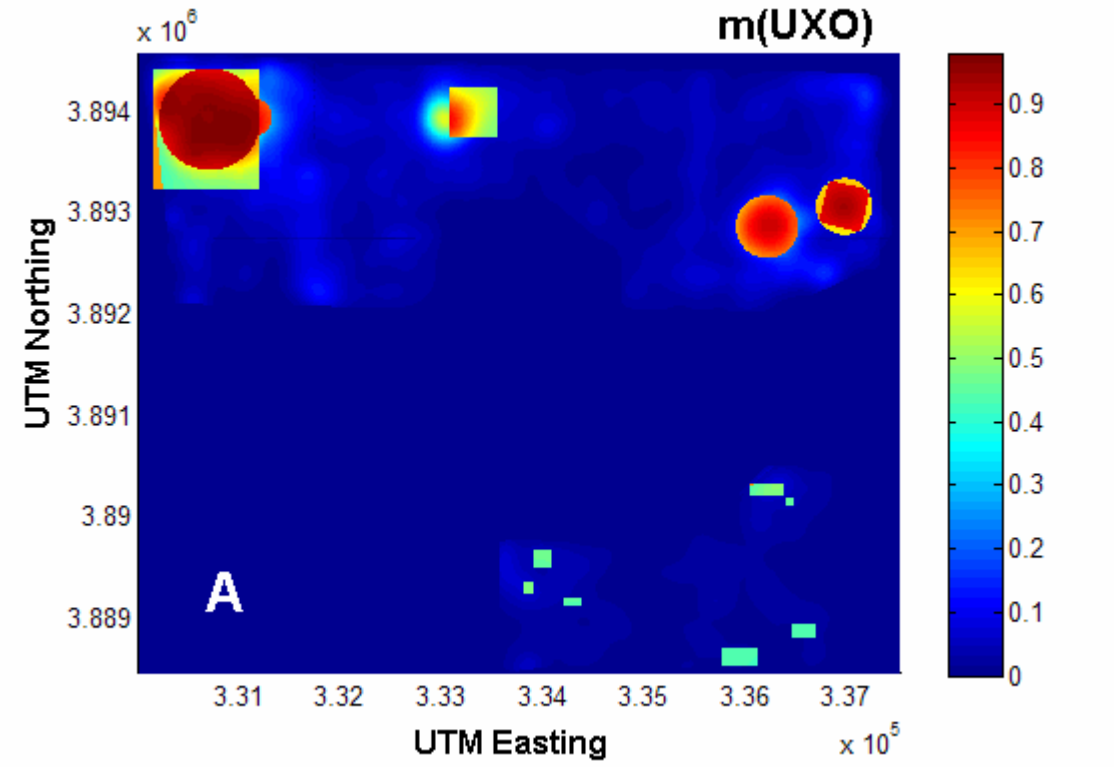
*Dempster-Shafer Approach.* In order to test the Dempster-Shafer data fusion algorithm implementation, a subset of feature layers were chosen for input and assigned preliminary probability mass functions, $m()$, for the UXO frame of discernment, $\{UXO\}$, $\{\neg UXO\}$, and $\{UXO, \neg UXO\}$. As detailed probability mass functions were unknown, probability masses were assigned to the extremal feature intensity values (0 and 1) for each layer. Probability mass assignments for intermediate feature intensity values were then generated through linear interpolation from the extremal values, as described earlier.

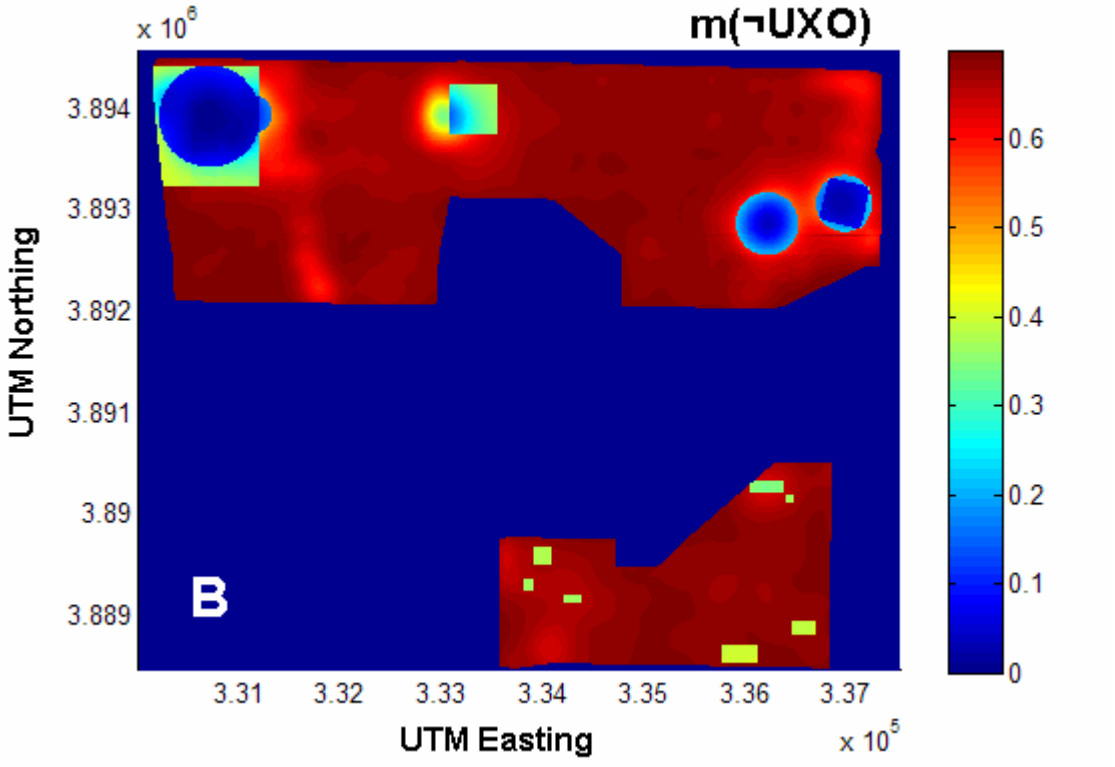
The assignments for the first set of selected feature layers are summarized in Figure 23. For zero crater density, all probability mass was assigned to uncertainty. This reflected the fact that an absence of craters provides neither positive nor negative evidence for the presence of UXO-related objects. At the maximum crater density observed (i.e., an intensity value of 1), a probability mass of 0.75 was assigned to the proposition that UXO was present and 0.25 to uncertainty, reflecting the assertion that relatively high crater densities are associated with the presence of UXO-related objects, although not unequivocally so. The probability mass assignments for each of the two manually-delineated feature layers followed a similar pattern. Where these features were not present, no information was provided, and thus the probability mass assignment went entirely to uncertainty. Where the features were present, an association with UXO was indicated by a non-zero assignment to $m(UXO)$ that was proportionate to the degree of belief that feature provided to the proposition that UXO was present. The remainder of the mass was assigned to uncertainty. Helicopter magnetometry uncertainty assignments reflected the fact that a lack of anomaly density provided partial evidence for an absence of UXO-related objects, while a high anomaly density provided partial evidence for the presence of UXO-related objects.

| | Feature Intensity | *m(UXO)* | *m(¬UXO)* | *m(UXO, ¬UXO)* |
|---|---|---|---|---|
| **Crater Density** | **0** | 0 | 0 | 1 |
| | **1** | 0.75 | 0 | 0.25 |
| **Helimag Anomaly Density** | **0** | 0 | 0.7 | 0.3 |
| | **1** | 0.7 | 0 | 0.3 |
| **Target Features** | **0** | 0 | 0 | 1 |
| | **1** | 0.8 | 0 | 0.2 |
| **Munitions Areas** | **0** | 0 | 0 | 1 |
| | **1** | 0.7 | 0 | 0.3 |

Figure 23. Probability mass assignments for Dempster-Shafer data fusion for extreme feature intensity values (0 and 1) of layers: helimag anomalies, manually detected craters, known bombing targets, and manually delineated munitions areas.

The probability mass assignments were input along with the corresponding feature layers to the Dempster-Shafer algorithm implementation. Figure 24 depicts the resulting output site-wide assessments of *m*(UXO), *m*(¬UXO), and *m*(UXO,¬UXO).


m(UXO)
x 10^6
3.894
3.893
3.892
3.891
3.89
3.889
UTM Northing
A
3.31
3.32
3.33
3.34
3.35
3.36
3.37
UTM Easting
x 10^5
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0


m(¬UXO)
x 10^6
3.894
3.893
3.892
3.891
3.89
3.889
UTM Northing
B
3.31
3.32
3.33
3.34
3.35
3.36
3.37
UTM Easting
x 10^5
0.6
0.5
0.4
0.3
0.2
0.1
0

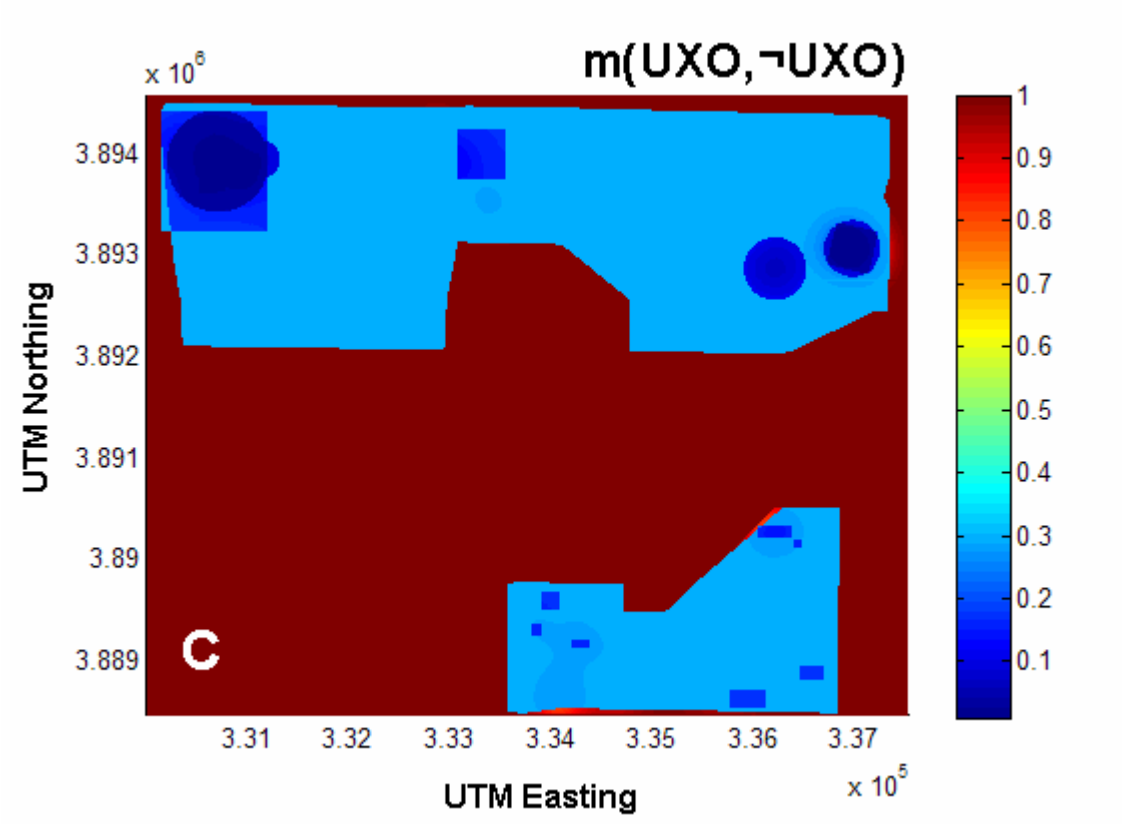



Figure 24. Dempster-Shafer data fusion output combining helimag anomaly density, manually detected crater density, known bombing targets, and manually delineated munitions areas at the Kirtland site.

Figure 24A is a color-scaled map depicting the output assessment of *m*(UXO), that is, the degree of belief the four lines of evidence assigned to the proposition that UXO was present. Dominating this map are the regions corresponding to known bombing targets and munitions areas, which generally coincide with relatively high densities of magnetic anomalies and craters. The potential munitions areas in the southern region of the Kirtland site exhibit a much smaller value assigned to *m*(UXO) in the data fusion output. The smaller value was due to the lack of corroborating evidence provided by the other feature layers. Figure 24B is a color-scaled map depicting the output assessment of *m*(¬UXO), or the proposition that UXO-related objects are not present. Since the magnetic anomaly feature layer was the only one providing evidence corroborating this proposition, we see high values of *m*(¬UXO) wherever magnetic anomaly density was low, and no other evidence contradicted it. Finally, Figure 24C is a color-scaled map depicting the output assessment of *m*(UXO,¬UXO), or uncertainty as to whether UXO is present or not. As expected, the data fusion algorithm assigned a high value to this parameter in areas where survey data were missing or unavailable. Elsewhere, mass assigned to uncertainty was much less, reflecting the increased certainty with which the presence or absence of UXO was supported by available evidence.

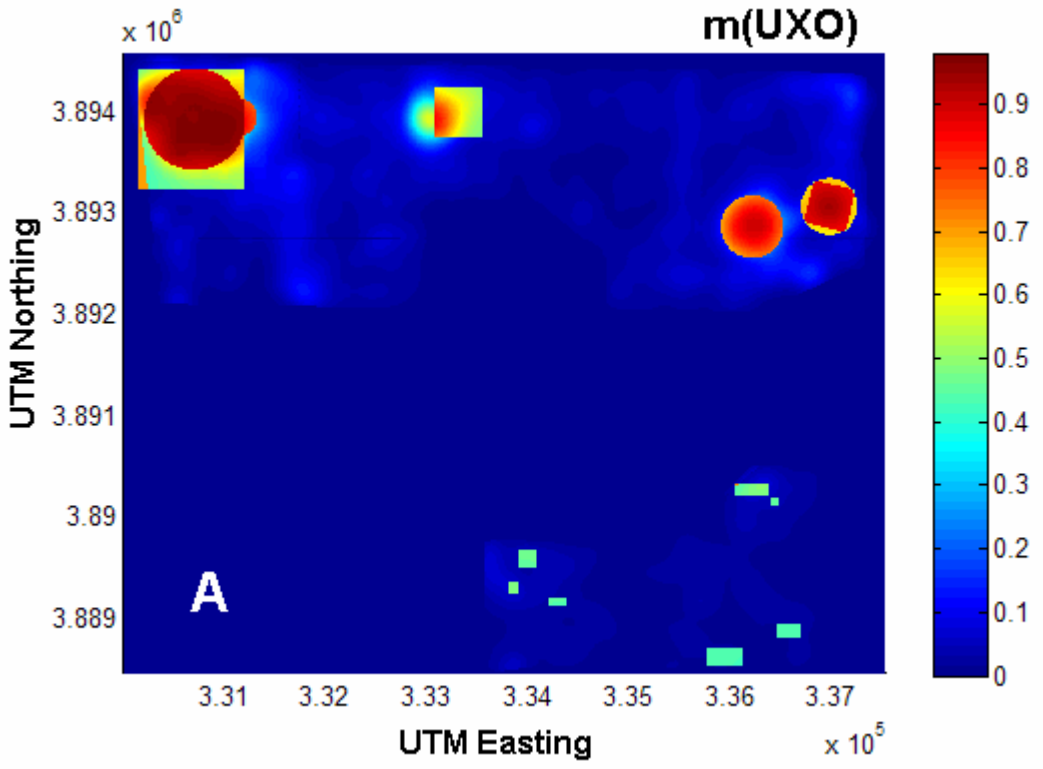
A comparison of Figures 18, 21, and 24 illustrates the difference between the data fusion approaches. Although each output depicts more or less the same areas of interest within the Kirtland site, the amount of information contained is very different. The weighted overlay shown in Figure 18, for instance, provides only limited indication as to the degree of UXO likelihood. The Bayesian output shown in Figure 21, however, indicates a relative propensity for or against the presence of UXO-related objects, with an output of 0.5 indicating complete uncertainty. Finally, the Dempster-Shafer output depicted in Figure 24 provides three output maps, collectively indicating the proportion of the available evidence that supports each of the three propositions of the frame of discernment at each point in the survey area.

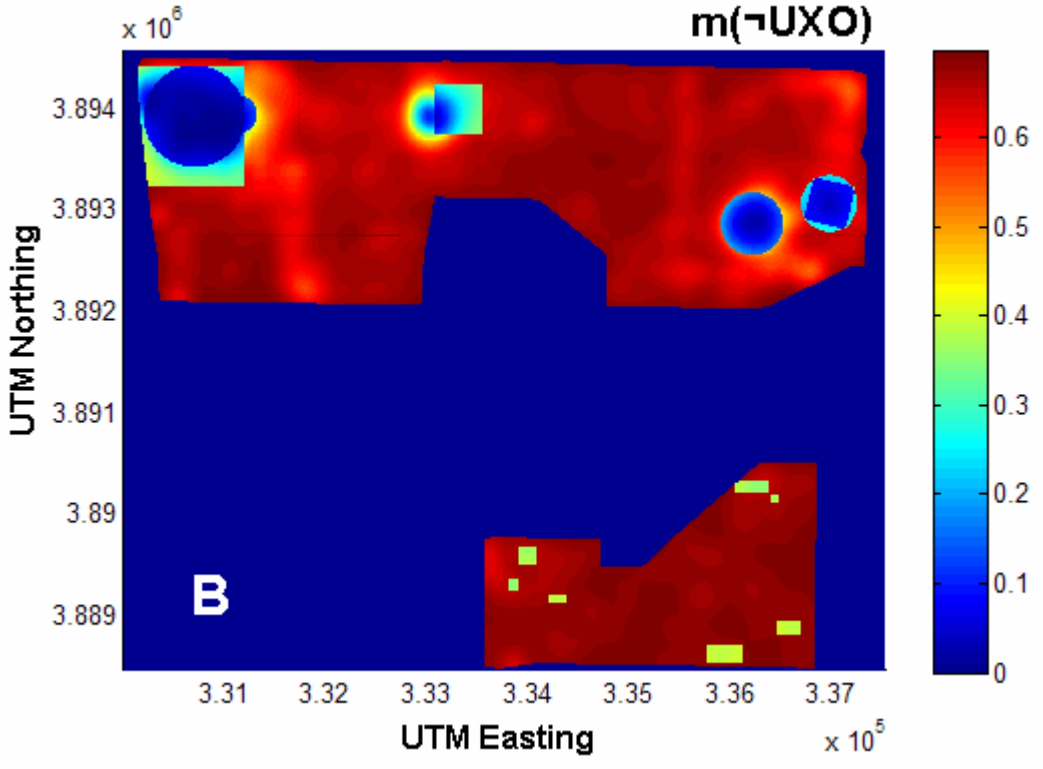
Next, an alternate scenario was tested where a different set of feature layers was utilized. In this case, it was simulated that manually detected craters and magnetic anomaly features were not available, and in their place, automatically detected craters derived from LiDAR data and threshold-applied, morphologically filtered, helimag signal were used. The probability mass assignments were adjusted accordingly to reflect an increased degree of uncertainty in their ability to correctly indicate UXO-related objects, as shown in Figure 25.

| | Feature Intensity | *m(UXO)* | *m(¬UXO)* | *m(UXO, ¬UXO)* |
|---|---|---|---|---|
| **Auto-detected Crater Density** | 0 | 0 | 0 | 1 |
| | 1 | 0.5 | 0 | 0.5 |
| **Threshold Helimag Density** | 0 | 0 | 0.7 | 0.3 |
| | 1 | 0.65 | 0 | 0.35 |
| **Target Features** | 0 | 0 | 0 | 1 |
| | 1 | 0.8 | 0 | 0.2 |
| **Munitions Areas** | 0 | 0 | 0 | 1 |
| | 1 | 0.7 | 0 | 0.3 |

Figure 25. Probability mass assignments for Dempster-Shafer data fusion for extreme feature intensity values (0 and 1) of layers: thresholded helimag signal with morphological filtering, automatically detected craters, known bombing targets, and manually delineated munitions areas.

Utilizing these assignments and the corresponding feature layers, the following output was obtained from the Dempster-Shafer based data fusion algorithm.


m(UXO)
x 10^6
3.894
3.893
3.892
3.891
3.89
3.889
UTM Northing
A
3.31
3.32
3.33
3.34
3.35
3.36
3.37
UTM Easting
x 10^5
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0


m(¬UXO)
x 10^6
3.894
3.893
3.892
3.891
3.89
3.889
UTM Northing
B
3.31
3.32
3.33
3.34
3.35
3.36
3.37
UTM Easting
x 10^5
0.6
0.5
0.4
0.3
0.2
0.1
0

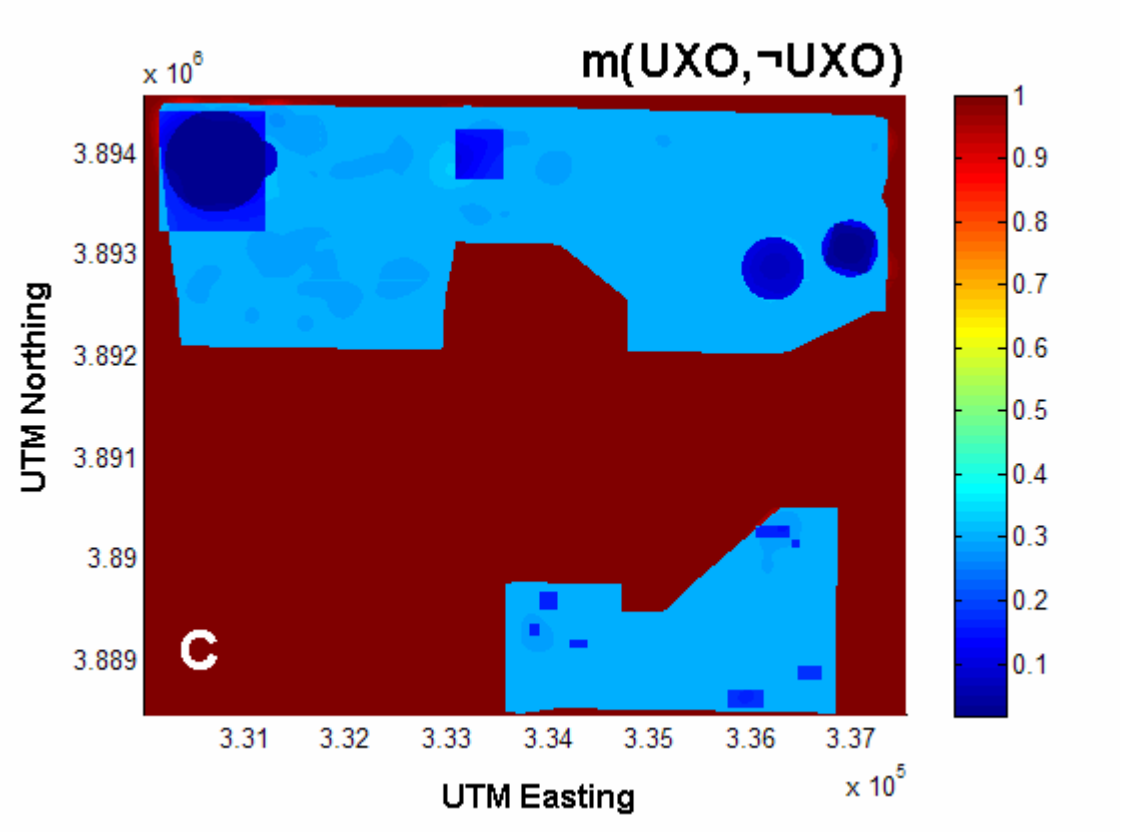



Figure 26. Dempster-Shafer data fusion output combining thresholded helimag signal with morphological filtering, automatically detected craters, known bombing targets, and manually delineated munitions areas at the Kirtland site.

As compared to the previous feature layer set, essentially the same output was obtained, although in Figure 26B slightly lower values were assigned to *m*(¬UXO) in many areas, reflecting the relatively higher densities observed throughout the site with this magnetometry feature layer.

A similar exercise was performed with an analogous set of feature layers available for the Pueblo site, utilizing the same probability mass assignment functions shown in Figures 23 and 25. In this example, magnetic anomaly density and automatically-detected crater location density feature layers were utilized. The results of Dempster-Shafer data fusion with these Pueblo feature layers are shown in Figure 27.

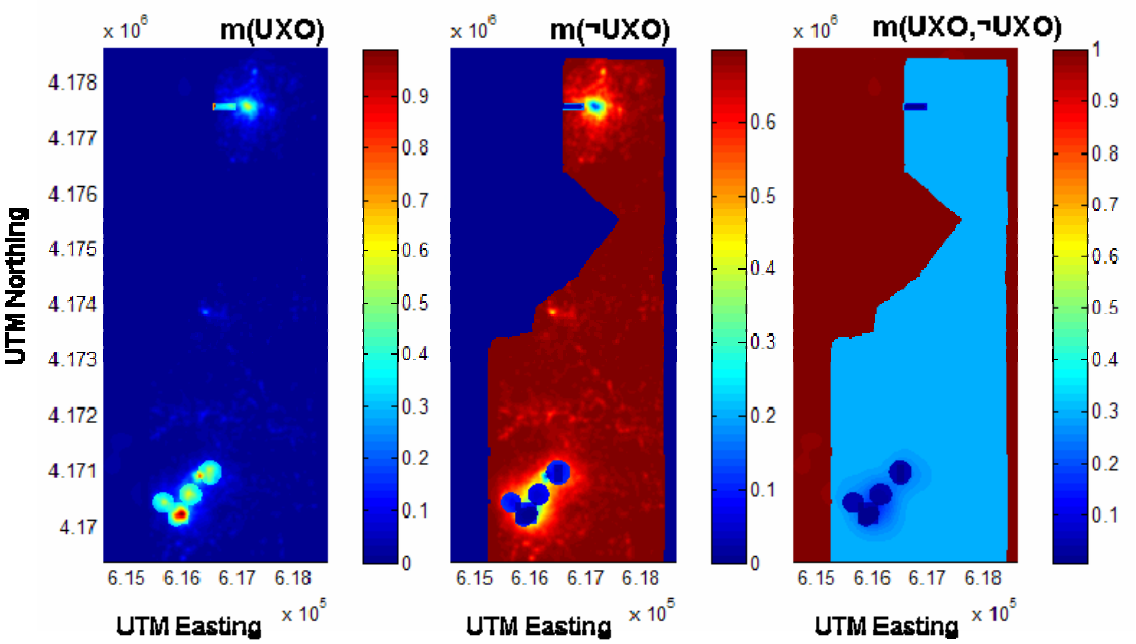



Figure 27. Dempster-Shafer data fusion output combining helimag anomaly density, automatically detected crater density, known bombing targets, and manually delineated munitions areas at the Pueblo site.

As with the Kirtland data fusion results, the highest values assigned to *m*(UXO) occurred where multiple lines of corroborating evidence supported the proposition that UXO-related objects were present, while the highest values assigned to *m*(UXO,¬UXO) occurred where no data were available. By setting a threshold of 0.3 for *m*(UXO), as shown in Figure 28A, a potential delineation of UXO-contaminated areas was generated. When overlaid with limited ground truth data available at the Pueblo site, as in Figure 28B, it can be seen that the delineated regions were consistent with ground truth.

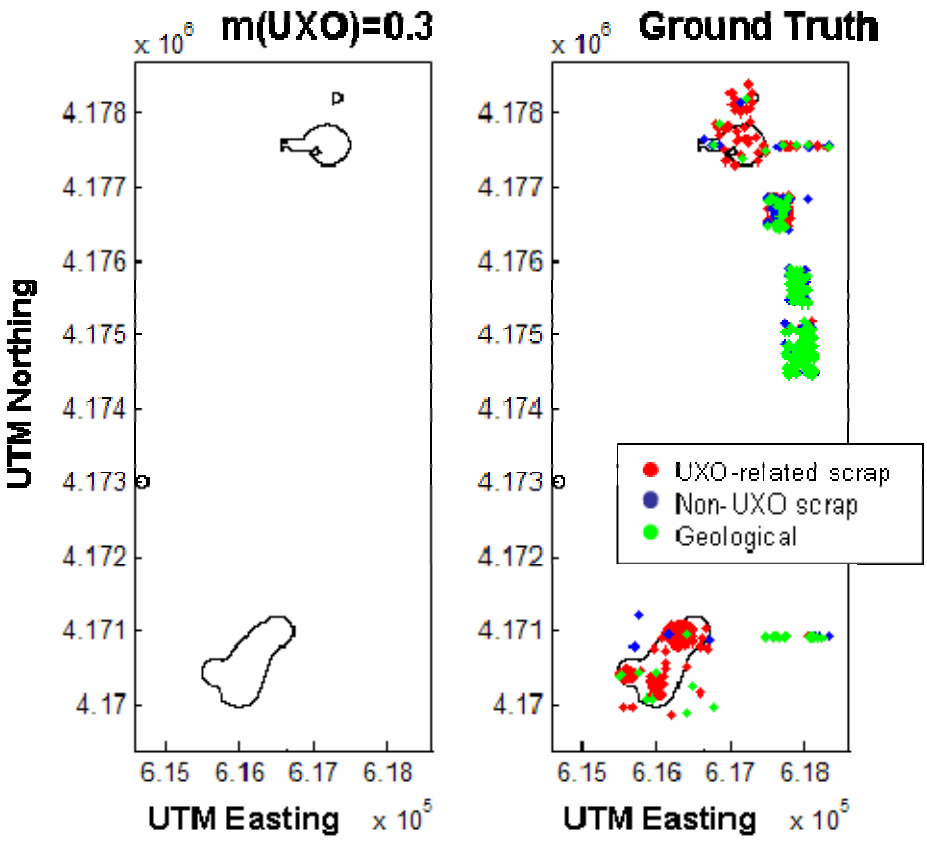
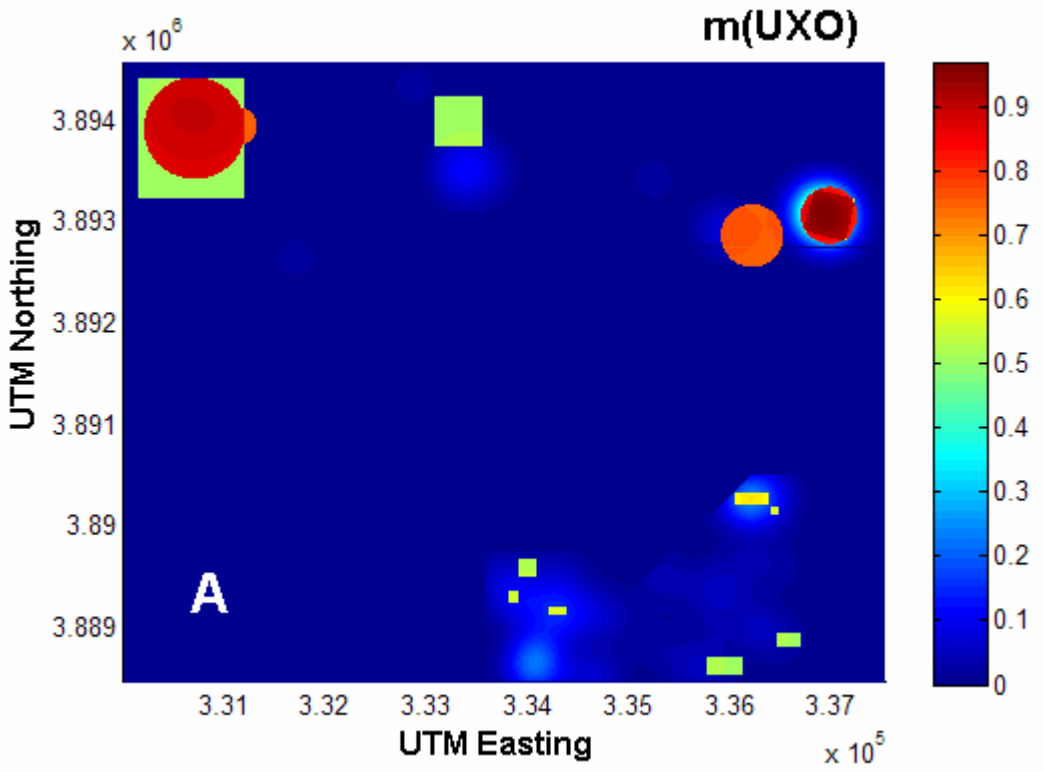
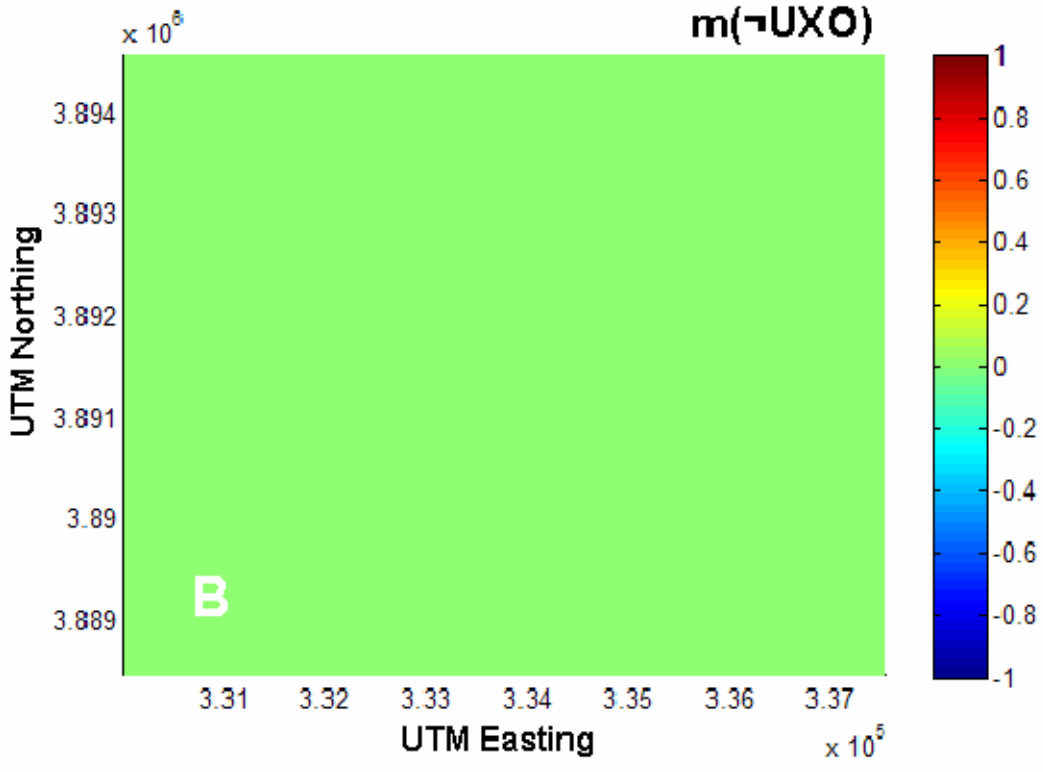



Figure 28. A preliminary delineation of areas of likely UXO contamination utilizing Dempster-Shafer data fusion output for the Pueblo site threshold at 0.3. Overlaid on this plot are limited truth data showing agreement with this assessment.

The effects of adding and subtracting feature layers were examined by recalculating the data fusion algorithm output with a reduced feature layer set. Figures 29 and 30 demonstrate the Dempster-Shafer algorithm's ability to function when input data streams are missing, but a comparison to Figure 24 shows the toll subtracting an information source can have on data fusion output. The effect of the removal of the magnetic anomaly density feature layer is shown in Figure 29, which depicts the output of Dempster-Shafer fusion of manually-located crater density with manually identified bombing targets and manually delineated potential munitions remediation areas.


m(UXO)
x 10^6
3.894
3.893
3.892
3.891
3.89
3.889
UTM Northing
A
3.31
3.32
3.33
3.34
3.35
3.36
3.37
UTM Easting
x 10^5
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0


m(¬UXO)
x 10^6
3.894
3.893
3.892
3.891
3.89
3.889
UTM Northing
B
3.31
3.32
3.33
3.34
3.35
3.36
3.37
UTM Easting
x 10^5
1
0.8
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
-0.8
-1

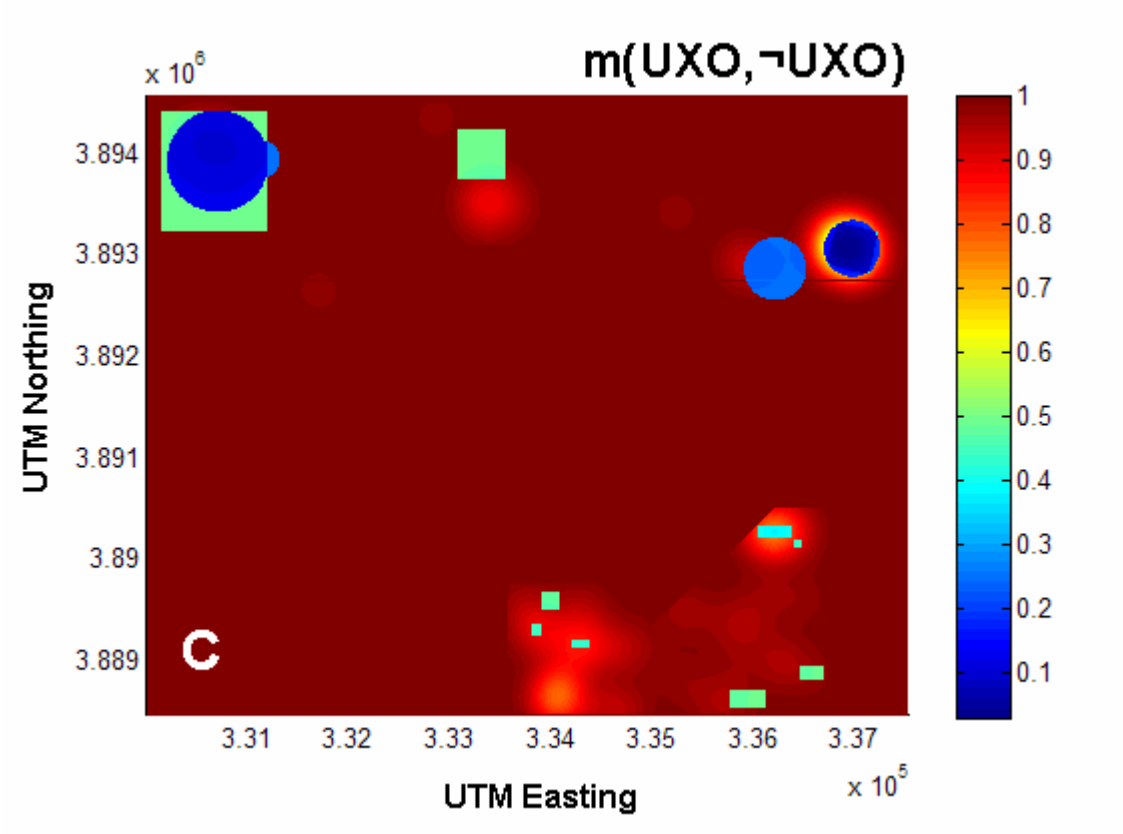
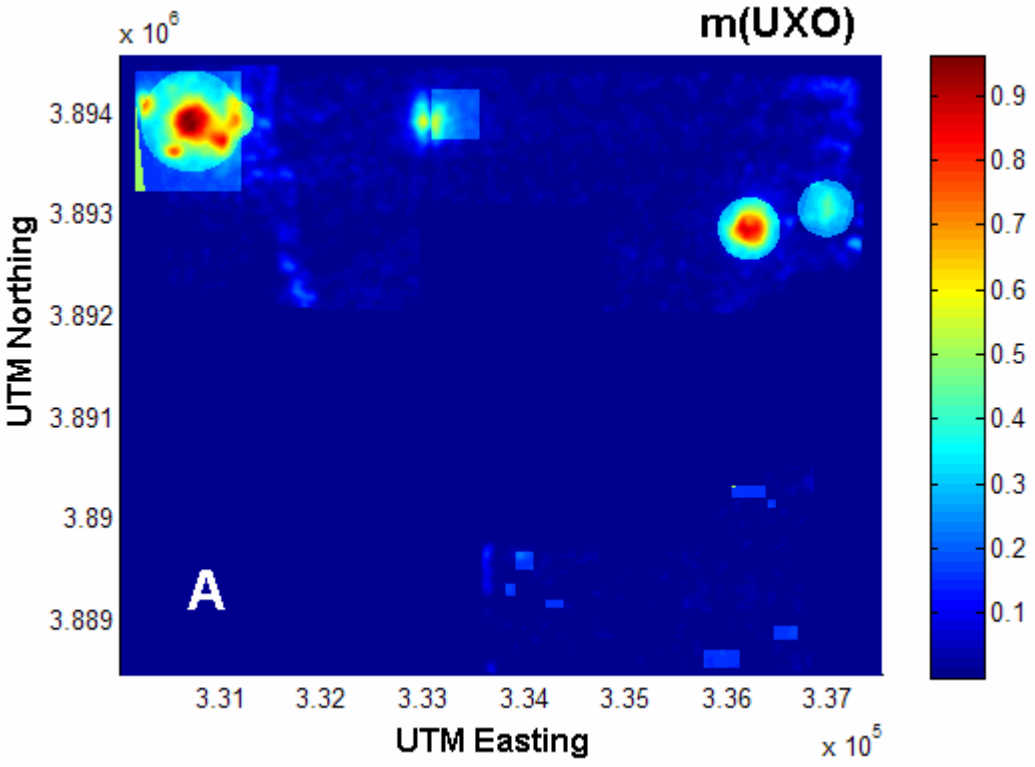
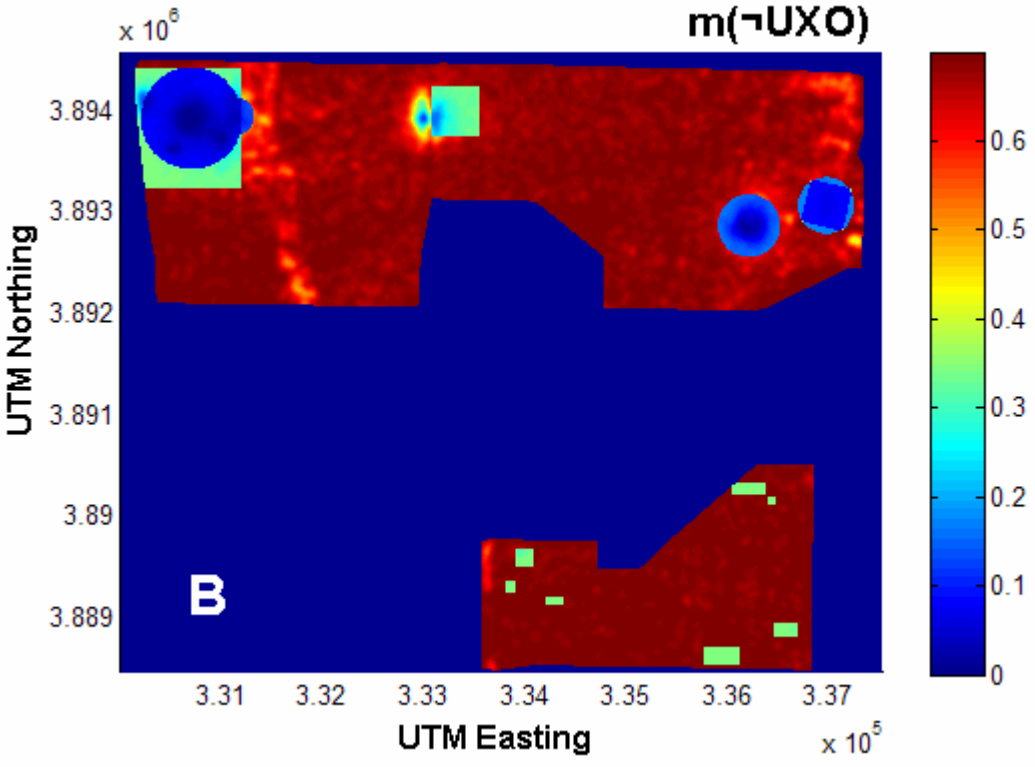



Figure 29. Dempster-Shafer data fusion output combining manually detected craters, known bombing targets, and manually delineated munitions areas.

The removal of the magnetic anomaly density feature layer resulted in a zero assignment of belief to *m*(¬UXO), as well as a relative increase in the belief assigned to *m*(UXO) for the areas delineated as potential munitions areas. In both cases, the changes were due to the absence of evidence that a lack of magnetic anomaly density provided to the proposition the UXO was not present.

Alternatively, removing the crater density feature layer resulted in the output shown in Figure 30, which depicts the output of Dempster-Shafer fusion of helicopter magnetometry density with manually identified bombing targets and manually delineated potential munitions remediation areas.


m(UXO)
x 10^6
3.894
3.893
3.892
3.891
3.89
3.889
UTM Northing
A
3.31
3.32
3.33
3.34
3.35
3.36
3.37
UTM Easting
x 10^5
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1


m(¬UXO)
x 10^6
3.894
3.893
3.892
3.891
3.89
3.889
UTM Northing
B
3.31
3.32
3.33
3.34
3.35
3.36
3.37
UTM Easting
x 10^5
0.6
0.5
0.4
0.3
0.2
0.1
0

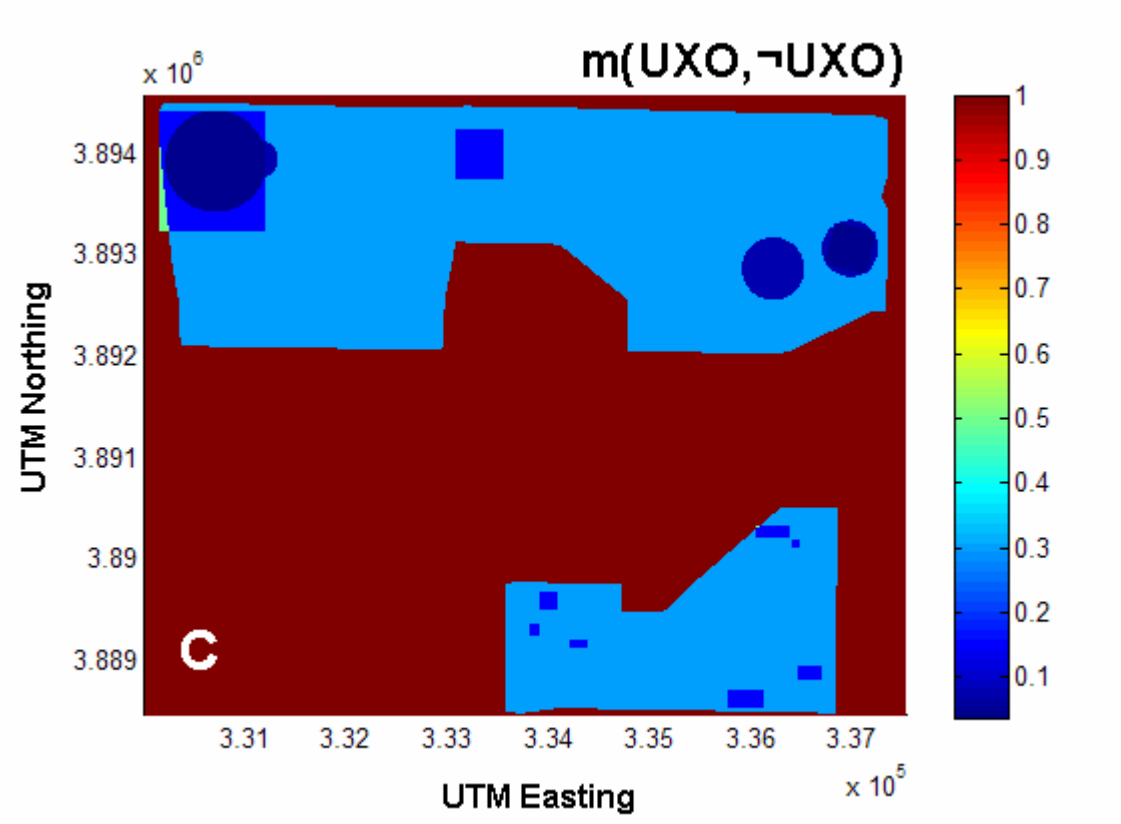



Figure 30. Dempster-Shafer data fusion output combining helimag anomalies, known bombing targets, and manually delineated munitions areas.

The removal of the crater density feature layer resulted in a relatively smaller assignment to *m*(UXO) for the manually-located bombing targets at the eastern and western edges of the Kirtland survey area, as these areas contained relatively high densities of manually identified craters.

*Hybrid Heuristic/Dempster-Shafer Approach.* As data fusion algorithm development progressed, it became clear that in certain instances heuristic rules were required to capture specific feature-to-feature interdependencies. For example, it is possible that manually delineated regions of interest in the feature set may describe locations or objects associated with significant non-UXO magnetic signal, such as fence lines, man-made structures, and pipelines. Accordingly, the impact of magnetic signals recorded at these locations on the output UXO assessment should be minimized or blocked. A second example reflects the incorporation of measures of data quality. The helicopter magnetometry data layer was estimated from point measurements that were acquired at non-regular spatial intervals and provided non-homogenous coverage of the survey site. Areas with sparse data point coverage in the magnetometry data provided less certain information than those with dense coverage. Knowledge of the quality of data point coverage needed to be reflected in the output UXO assessment. Thus, in a general sense it became important to explore means by which these kinds of knowledge could be

incorporated into the data fusion framework. The following implementations were developed for the two heuristic rules just described.

A simple rule incorporating the concept of blocking known non-UXO magnetic signals was implemented as follows. First, feature layers associated with significant non-UXO magnetic signal were flagged as such. Second, feature layers susceptible to this form of magnetic interference were flagged as such. Finally, a general rule was added to the data fusion algorithm indicating that whenever a feature layer belonging to the latter set was being used to update a current UXO assessment, the regions of that feature layer coinciding with those identified in the former (blocking) set were shrunk to a value of zero.

The effect of this heuristic rule can be observed in the data fusion output generated with and without the rule in place, as shown in Figure 31.

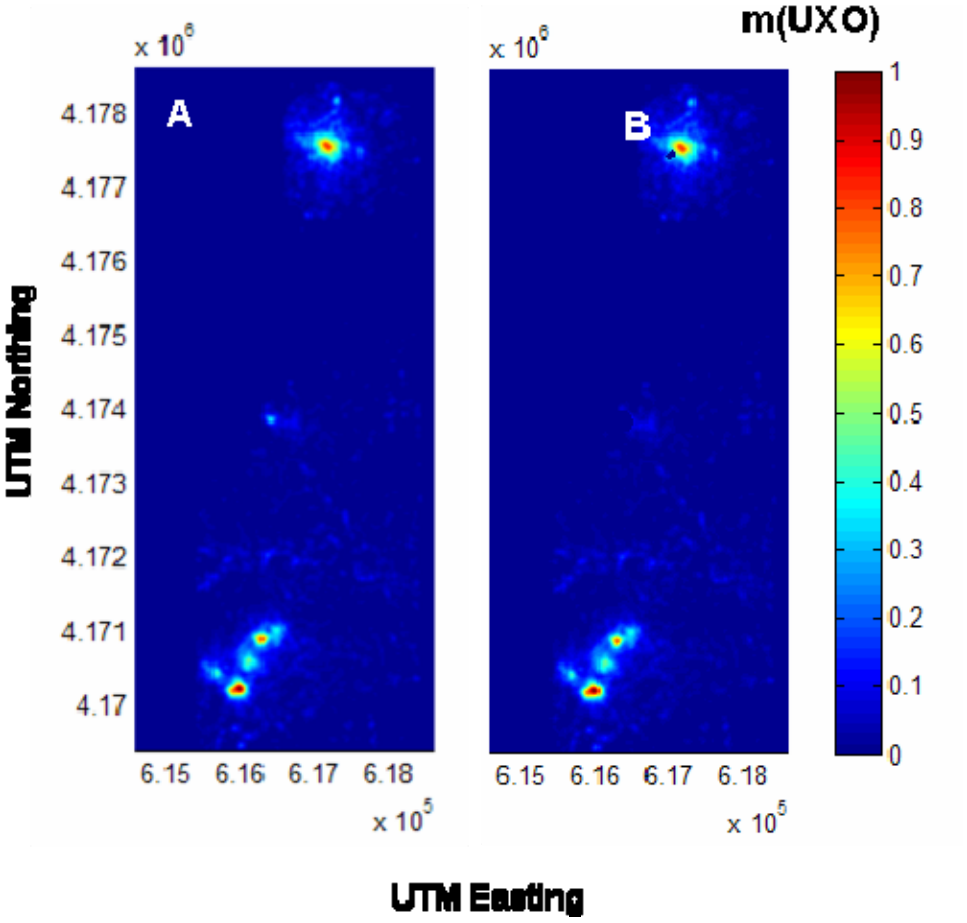



Figure 31. UXO assessment before and after incorporation of a heuristic rule blocking magnetic signal in regions containing known interferences.

The principal difference lies in the reduction of belief assigned to *m*(UXO) in an area in the center of the survey region, near the point (4.174, 6.16). This area, which was manually identified as containing man-made structures, resulted in a relatively high density of magnetic anomalies that were presumably not due to UXO-related objects and an artificially high assignment of probability mass to the region. After application of the heuristic rule, the feature density from magnetic anomalies in this region was no longer considered in the final data fusion output.

Implementation of a rule incorporating data density was accomplished as follows. The first step was the development of a density metric for the helicopter magnetometry data feature layer. As part of the year one effort, a layer containing data point density was calculated as the number of data points observed within a square meter on 1m common grid. A metric was derived from this value as $1/(1+\text{density})^2$ and ranged in value from 1 (corresponding to zero coverage) to close to 0 as coverage increased. The resulting map of density metric values was then convolved with a Gaussian kernel for smoothing to allow for neighborhood effects. The final data density metric is shown in Figure 32.

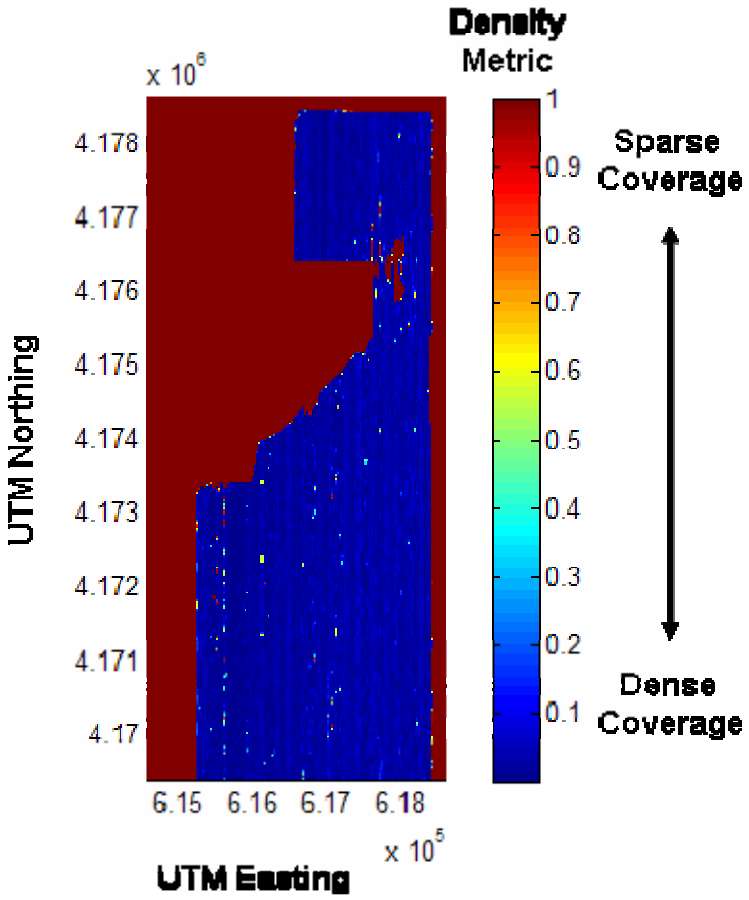



Figure 32. Site-wide values for the helicopter magnetometry data density metric.

Once calculated, the data density metric was used to modify probability mass assignments for magnetometry feature layers. This was accomplished by, again flagging the set of feature layers susceptible to this rule. When such a feature layer was being used to update an output UXO assessment, the data density metric layer was loaded and used to modify the probability mass assignments as follows: grid points with a metric of one were assigned a probability mass associated with complete uncertainty. Grid points associated with a density metric close to zero were assigned the probability masses indicated by the appropriate probability mass functions for that layer. Finally, grid points with a data density metric intermediate between one and zero were assigned probability masses equal to a linear combination of these two extremes, scaled according to the value of the density metric. The effect of this heuristic rule can be observed in the data fusion output generated with and without the rule in place, as shown in Figure 33.

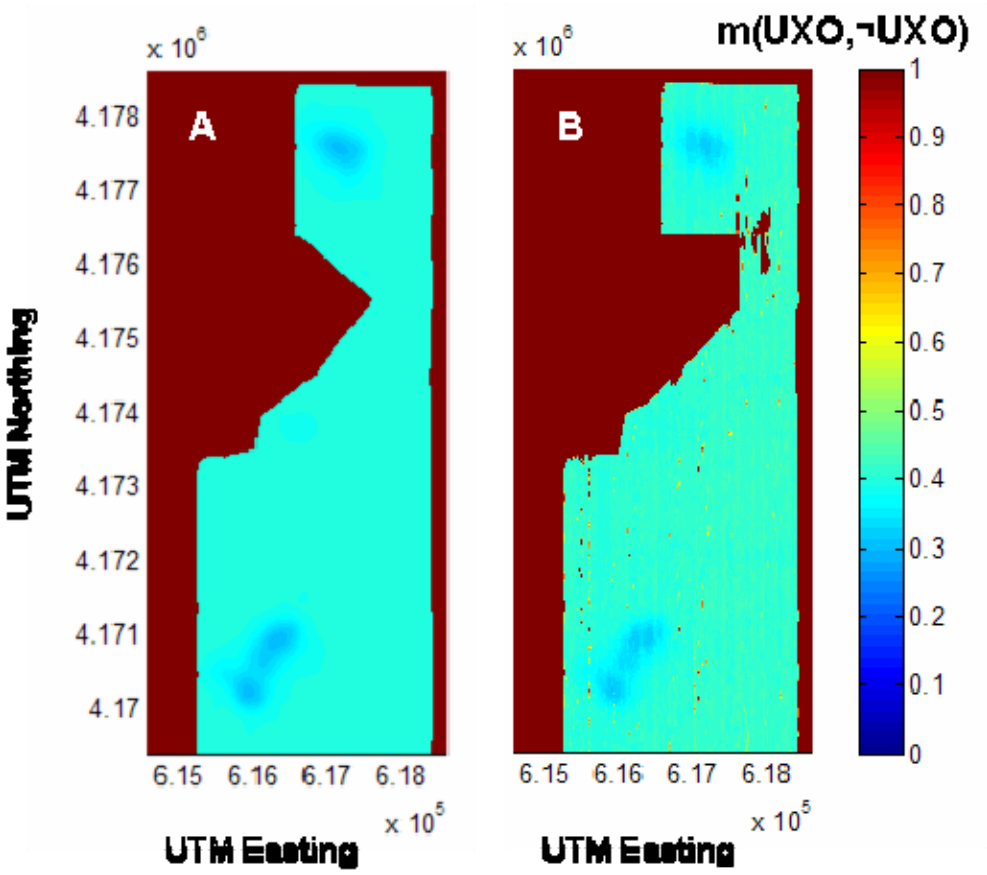



Figure 33. UXO assessment (A) before and (B) after incorporation of a heuristic rule incorporating a data density metric for helicopter magnetometry feature layers.

The impact of the heuristic can be seen in the increase in assignment to *m*(UXO,¬UXO), which is most pronounced in areas with zero data density where the assignment of belief goes almost completely to uncertainty. A revised data fusion output for the Pueblo site,

incorporating both of the developed and implemented heuristics rules is shown in Figure 34.

Heuristic rules such as these allow for the incorporation of specific knowledge regarding relationships between feature sets. In the first example, the incorporation of a rule blocking magnetometry signal in regions with man-made structures led to a reduction of false positive indications of UXO. In the second example, the incorporation of a rule attenuating the impact of magnetometry signal in areas of sparse data density allowed for a more accurate assessment of the uncertainty present in the resulting Dempster-Shafer data fusion output maps.

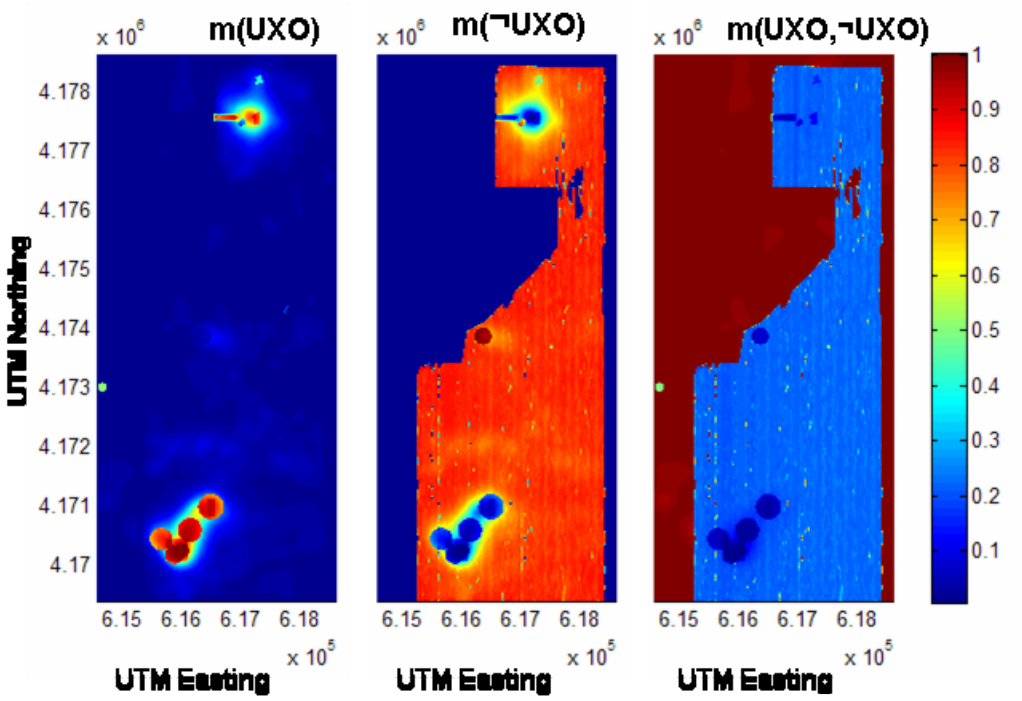



Figure 34. Dempster-Shafer data fusion output combining thresholded helimag signal with morphological filtering, automatically detected craters, known bombing targets, and manually delineated munitions areas at the Pueblo site. Heuristic rules involving feature blocking and data point density metrics are implemented.

**Summary of Results from Year Two Tasks**

The major accomplishments for Year Two are listed below:

- A wavelet filtering algorithm was developed to improve the automatic crater identification algorithm. Filtering reduced false positive crater identifications by removing surface texture artifacts.
- A statistical nearest-neighborhood clustering filter algorithm was developed to improve the magnetic anomaly data by locating anomalies that were part of non-uniform groupings. The clustering algorithm reduced background noise and false positive assessments of UXO in the final data fusion output.
- The development of a Data Fusion Framework prototype based on a hybrid Dempster-Shafer theory.
- A methodology for the input and registration of disparate data and feature streams from wide-area sensing technologies and other site-specific knowledge.
- A methodology allowing the input of meta-information regarding the relationship between these lines of evidence and the presence (or absence) of UXO or UXO-related objects.
- A method by which heuristic rules can be incorporated into the Data Fusion Framework in order to take advantage of specific known interdependencies between feature layers.
- Demonstration of the Data Fusion Framework prototype's ability to successfully provide useful output assessment of UXO likelihood from an assortment of WAA assessment data and features.

Project MM-1510 leveraged data and feature sets acquired in the ESTCP Wide-Area Assessment Pilot Program. In year two, data were obtained from SERDP that had been acquired at the Kirtland ESTCP WAA pilot program survey site in New Mexico. While some effort was directed to improved feature selection, the majority of the year's effort was devoted to the development of a data fusion framework.

Algorithms for the automatic identification of craters and magnetic anomalies developed in year one were improved and optimized. These optimized algorithms were applied to data from both the Pueblo, CO and the Kirtland, NM data sets to generate feature inputs for data fusion. A crater detection algorithm, developed utilizing LiDAR data from the Pueblo site, was shown to be effective when applied to LiDAR data from the Kirtland site. A wavelet filtering algorithm was developed and utilized to reduce apparent false positive crater identifications by removing surface texture artifacts.

Magnetic anomaly data were enhanced through the development of a statistical nearest-neighborhood clustering filter algorithm that proved effective in locating anomalies that were part of non-uniform groupings. As UXO-related objects tend to be clustered around bombing targets and munitions ranges, this algorithm had the impact of selectively removing non-UXO related magnetic anomalies from the feature set and reducing false positive assessments of UXO in the final data fusion output.

Development of the data fusion framework centered on generating methodologies for three component tasks: generation of feature layers from input data and feature sets, formulation of a structured means allowing the input of meta-information regarding the relationship between feature layers and the presence (or absence) of UXO-related objects, and development of an algorithmic means for combining the evidence presented by each available feature set and corresponding meta-information. A methodology for the input and registration of disparate data and feature streams from wide-area sensing technologies and other site-specific knowledge was developed. The input methodology was based on expanding the approaches taken in year one, which involved the generation of site-wide feature intensity maps for each input feature stream. Three distinct types of feature streams were identified: point-located features and manually delineated regions, both of which were binary-valued, and continuously valued, intensity-based features. For each feature type, methods were specified for the construction of appropriate feature maps, or layers, suitable for data fusion. A structured methodology allowing the input of meta-information regarding the relationship between these lines of evidence and the presence (or absence) of UXO or UXO-related objects was created. This methodology provides users with an intuitive means to define a functional relationship between feature intensity values and the likelihood or degree of belief conferred upon the hypotheses under consideration.

Potential algorithmic means for combining the evidence presented by input of wide area UXO assessments feature sets and their corresponding meta-information were examined. This work focused on Bayesian and Dempster-Shafer theory based approaches to data fusion. Prototype software implementations of data fusion frameworks with these approaches were developed utilizing the MATLAB computational platform and evaluated with the input feature sets generated in years one and two of the project. Bayesian-based data fusion was shown to be inherently problematic for data fusion of WAA data as it required the specification of a functional dependence that was ill-formed for features such as craters and manually identified target and munitions areas. A Dempster-Shafer based algorithmic approach to data fusion was shown to be adept at handling both features with and without well-defined functional dependence with UXO-related objects.

The prototype Dempster-Shafer algorithm was evaluated using various combinations of input feature layers, including reduced feature sets. While reduced feature layer inputs provided realistic assessments, the incorporation of additional feature layers improved the output assessment. This evaluation demonstrated the stability and robustness of the Dempster-Shafer algorithm, as well as the complementary nature of the information provided by the input feature layers. Preliminary results obtained with the prototype Dempster-Shafer based data fusion framework were shown to agree well with ground truth data available at the Pueblo site. The prototype framework was further refined by the incorporation of two heuristic rules to accommodate specific a priori knowledge about input data and feature layers.

# Conclusions

A generalized method for processing input data feature streams from UXO WAA survey efforts was developed. The method requires the generation of a corresponding feature intensity map and the specification of a functional relationship between a feature's intensity and the hypotheses (i.e., the presence or absence of UXO as well as the possibility that at certain values or in certain areas the feature doesn't support a determination either way) that are supported by the feature's intensity values. The development of this approach is significant, as it requires only a limited number of specifications to be imposed on these two inputs, allowing a wide range of feature sets and relationships to be formatted and input for data fusion. This flexibility is crucial, as the disparate nature of the data and features available from potential WAA survey techniques presents a significant impediment towards adoption of more basic data fusion approaches. As part of this work, various feature sets derived from both ESTCP performers and from customized feature extraction algorithms developed at NRL were utilized as inputs to data fusion, demonstrating the flexibility of the approach.

Successful feature extraction algorithms developed in year one were further optimized and validated against new data, including an automatic crater detection algorithm and a nearest-neighbor clustering filter to selectively remove false-positive magnetometry anomalies. Heuristic, Bayesian, and Dempster-Shafer theoretic algorithms for combining evidence presented in feature layers were investigated as possible engines for a UXO WAA data fusion framework prototype. These were implemented as MATLAB code and evaluated with feature layers generated from both the Pueblo and Kirtland site data acquired by performers in the ESTCP WAAPP. The Dempster-Shafer approach, with its ability to quantify uncertainty about evidence, was shown to be the most appropriate approach for the UXO problem and proved to be the most successful of the three. The ability to incorporate heuristic rules regarding specific dependencies between input feature layers into the Dempster-Shafer based data fusion framework prototype was described and demonstrated utilizing two specific examples. The first demonstrated a reduction of false positive indications of UXO by utilizing a feature layer comprised of manually identified man-made structures to selectively block magnetometry-derived features. The second demonstrated an adjustment of the impact of magnetometry-derived features on the output assessment of UXO to accurately reflect the uncertainty associated with increased magnetometry data sparseness in some areas of the helimag survey. The prototype data fusion framework developed was able to delineate areas of likely contamination while providing reasonable estimates of the likelihood of that contamination given supporting observational evidence and a priori knowledge. Preliminary results were compared with limited ground truth data available at the Pueblo site and agreed well.

While successful data fusion requires complementary data sets for input, the key theoretical advantage for wide-area assessment is the ability to reduce false positives while retaining high detection rates. The framework described is flexible, tolerating missing data and allowing multiple configurations of potential input data streams, as well

as scalable, allowing new data streams to easily be included in the assessment. Further, the impact of available and new data streams on the output can be readily quantified. When presented with reduced sets of input feature layers, the prototype provided reasonable, although less accurate, assessments of UXO contamination, demonstrating both the robustness of the approach and the improvement provided by data fusion of feature layers containing complementary information. One challenge is that the structured input methodology requires the specification of each feature layer's relationship to the presence or absence of UXO. However, the input methodology allows specification to be accomplished in a highly flexible manner. The user has the ability to input specifications that vary from simple, intuitive estimations based on expert knowledge to detailed functional relationships based on empirical evidence of sensor performance. Thus, the data fusion framework is capable of utilizing all the information and observation evidence available, without necessarily requiring that the exact same inputs be present for assessment. This flexibility is an important feature of the data fusion approach as it is expected that, for a number of reasons, it will rarely be the case that exactly the same types or quality of data will be available for analysis each time a wide-area UXO assessment is performed.

Future work is directed towards the development, evaluation, and optimization of a prototype Dempster-Shafer based data fusion framework. In the final year, a prototype software implementation of the data fusion framework will be developed that is suitable for demonstration and evaluation by SERDP at various sites. In addition, the development of a final production-grade data fusion framework that is well-suited for independent use by site administrators will be planned in close coordination with SERDP and potential end-users, and in partnership with an existing GIS software vendor. It is expected that the development and demonstration of the final production-grade data fusion framework will be completed under a follow-on ESTCP program.

# References


1. Available at URLs <http://www.estcp.org/ux/#Wide> and http://www.serdp.org/research/UXO.html
2. Luo, R.C., Yih, C.C., Su, K.L., "Multisensor Fusion and Integration: Approaches, Applications, and Future Research Directions," *IEEE Sensors Journal* **2**(2) 107-119 (2002)
3. Luo, R.C., Su, K.L., "A Review of High-level Multisensor Fusion: Approaches and applications," *Proc. of IEEE Intl. Conf. on Multisensor Fusion and Integration for Intelligent Systems*, Taipei, Taiwan, R.O.C., Aug., 1999.
4. B.J. Johnson, T.G. Moore, B.J. Blejer, C.F. Lee, T.P. Opar, S. Ayasli, and C.A. Primmerman, "A Research and Development Strategy for Unexploded Ordnance Sensing," SERDP project UX-860 final report, April 1996. Available at http://www.serdp.org/research/UXO.html
5. The Mathworks, Inc., "MATLAB Numerical Analysis Software Suite," <http://www.mathworks.com>.
6. Available at URL < http://www.estcp.org/links/WAA-Pilot-Video.cfm>
7. Yan Zhang; Collins, L.M.; Carin, L., "Unexploded ordnance detection using Bayesian physics-based data fusion," *Integrated Computer-Aided Engineering* **10**(3), 231-47, (2003)
8. Collins, L.M.; Zhang, Y.; Carin, L., "Model-based statistical sensor fusion for unexploded ordnance detection," *Proceedings of IEEE International Geoscience and Remote Sensing Symposium. IGARSS 2002*, 24-28 June 2002, Toronto, Ont., Canada, vol. 3, pg. 1556-9
9. I. Shamatava, F. Shubitidze, c.c. Chem, H.S. Youn, K. O'Neil, K. Sun, "Potential benefits of combining EMI and GPR for enhanced UXO discrimination at highly contaminated sites," Proceedings of SPIE Vol. 5415, 1201-1210, (2004)
10. L.M. Collins, Y. Zhang, J. Li, H. Wang, L. Carin, S.J. Hart, S.L. Rose-Pehrsson, H.H. Nelson, and J.R. McDonald, "A Comparison of the performance of Statistical and Fuzzy Algorithms for Unexpoded Ordnance Detection," IEEE Transactions on Fuzzy Systems, **9**(1), 17-30,( 2001)
11. "Intelligent Data Fusion for Wide-Area Assessment of UXO Contamination. SERDP Project MM-1510. 2006 Annual Report" Rose-Pehrsson, S.L., Johnson, K.J., Minor, C.P, Guthrie, V.N. NRL Memorandum Report, NRL/MR/6180—07-9039, April 20, 2007.
12. KDE Toolbox for MATLAB authored by Alexander Ihler and available at URL < http://www.mathworks.com/matlabcentral/fileexchange/loadFile.do?objectId=7800&objectType=File>
13. S. James Press, "Bayesian Statistics: Principles, Models, and Applications," John Wiley & Sons, New York 1989.
14. D.S. Sivia, "Data Analysis: A Bayesian Tutorial," Oxford University Press Inc., New York 1996
15. G. Shafer, "A Mathematical Theory of Evidence, " Princeton University Press, Princeton, NJ 1976
16. Lawrence A. Klein, "Sensor and Data Fusion: A Tool for Information Assessment and Decision Making," SPIE Press, Bellingham, WA 2004.
17. Kari Sentz and Scott Ferson, "Combination of Evidence in Dempster-Shafer Theory," Sandia National Laboratories Report SAND 2002-0835, April, 2002.
18. Z. Yi, Y. Khing, C.C. Seng, Z.X. Wei, "Multi-ultrasonic sensor fusion for autonomous mobile robots," *Proceedings of SPIE*, Vol. 4051, 314-321, (2000)

19. A. Sarkar, A. Banerjee, N. Banerjee, S. Brahma, B Kartikeyan, M. Chakraborty, K.L. Majumder, "Landcover classification in MRF context using Dempster-Shafer fusion for multisensor imagery," *IEEE transactions on image processing* **14**(5), 634-45 (2005)

20. M. Raza, I. Gondal, D. Green, R.L. Coppel, "Fusion of FNA-cytology and gene-expression data using Dempster-Shafer Theory of evidence to predict breast cancer tumors," *Bioinformation* **1**(5), 170-5, (2006)

21. Dr. Herb H. Nelson, personal communication (2007)